Intersubjectivity, Empathy and Nonverbal Interaction

Nicola Jane Plant

n.j.plant@qmul.ac.uk

School of Electronic Engineering and Computer Science

Queen Mary University of London

Submitted in fulfilment of the requirements of the Degree of Doctor of Philosophy

May 2017

Declaration

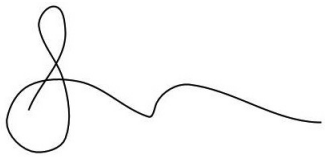
I, Nicola Jane Plant, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:



Date: 6th May 2017

Abstract

Empathy is thought to involve cognitive processes that depend on the simulation of another's experiences. Embodiment has a key role for empathy as vehicle for recreating the experience of another. This thesis explores the validity of this claim by investigating what people do when *communicating* about their experiences. In particular, what is the contribution of our embodied resources such as gestures, postures and expressions to empathy and intersubjectivity?

These questions are explored against two corpora of dyadic interactions. One features conversations of people describing recalled embodied experiences to each other, such as painful or pleasant bodily experiences like a headache or laughing. The other features a series of interactions designed to emulate informal conversations. The analysis uses hand coded gestures, feedback and clarification questions, body movement data and a new approach to quantifying posture congruence. The analysis shows the embodied responses observed within these interactions are intentionally placed and formulated to facilitate the incremental process of a conversation as a joint activity. This is inconsistent with accounts that propose there is an automatic and non-conscious propensity for people to mimic each other in social interactions.

Quantitative analysis show that patterns of gesture type and use, feedback form and posture differ systematically between interlocutors. Additionally, results show that resources provided by embodiment are allocated strategically. Nonverbal contributions increase in frequency and adjust their form responding to problems in conversation such as during clarification questions and repair. Detailed qualitative analysis shows the instances that appear to display mimicry within the interaction function rather as embodied adaptations or paraphrases. In their contrast with the original contribution they demonstrate a specific understanding of the type of experience being conveyed. This work shows that embodiment is an important resource for intersubjectivity and embodied communication is specifically constructed to aid the collaborative, sequential and intersubjective progression of dialogue.

Acknowledgements

This work is supported by the Media and Arts Technology programme, EPSRC Doctoral Training Centre EP/G03723X/1

Firstly, I would like to thank my supervisor Professor Patrick Healey not only for his guidance and intuition, but also for being extraordinary patient and never telling me off for being slow, giving support that I would get there in the end. I would next like to thanks my internal committee, firstly my secondary supervisor Dr Martin Welton, who gave me resources that inspired and expanded my vision early on. Secondly, Dr Graham White who was always gently encouraging and placed everything together for me while I had my head in a cloud of phenomenology. I would like to thank the thinking writing team at QMUL for organising the most pleasant and productive writing retreats and bootcamps that helped me get back on track when I'd lost confidence when writing up. I would also like to extend this to the rest of the academic staff in the Cognitive Science Research Group, for taking the time to suggest methods or ideas about my research as well always doing interesting research. I would like to thank Richard Kelly, the programme organiser at the time, for making things so easy and fun for me throughout the process. I would like to give a huge thanks to the rest of the Media and Arts Technology Pioneers, the later cohorts and PhD students of the Cognitive Science Research Group for always being around to hang out, inter-annotate, proof read, have a laugh and share things with. Lastly, I turn to non-academic support. My friends, for always being excited for me and putting up with me going AWOL for two and a half years. My family, always being so proud that I would finally make it, their kindness, generosity and positivity over the whole process.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Embodiment and the Other

To be embodied, the natural condition of human life, is to inhabit a form. For each of us our own embodiment is characterised experientially through the sensations of the body and kinaesthetically from a set of movement possibilities provided through our body. Typically, the Self[1] is characterised by the unequivocal access to our own embodiment through the immediate experience of sensations. If your head hurts, only you feel the pain sensation associated with the experience, and that is a direct experience of your embodiment. The embodiment of the Other[2] is perhaps a more difficult definition; it is differentiated from the Self by an indirect experience of the Other's body - an experience mediated by sense perception of the Other's animate exterior, their outward living body. Without direct access to inner experiences like we have to the Self, how do we understand the experience of the Other?[3] This understanding is widely known as Empathy. An empathetic experience is typically defined as a vicarious understanding the affective state (for instance; emotionally, attitudinally etc) of another. The term empathy refers to this capability but the capacity in which empathy is experienced is discussed from many different viewpoints in different disciplines. Empathy is conceptually distinct from Sympathy, although these terms are often used interchangeably. Sympathy does not refer to understanding

[1] The 'Self' as referred to in this context (with a capital S) is recognised as our own living consciousness.

[2] The 'Other' as referred to in this context (with a capital O) is recognised as another living consciousness.

[3] Of course, there is the challenge of how do we know that the Other is a living body to begin with? However, this work addresses its question by already assuming the existence of a living Other.

the experience of another, but the experience of expression of compassion or commiseration in response to another's experience or situation. Sympathy often culminates in offers of support that extend beyond demonstrations of empathetic understanding.

1.2 Intersubjective experiences

The key problem of concern here is not just how we understand the 'Other' but how we can have a mutual or *intersubjective* understanding of each other's experience? An assumption that: I know you have an understanding of my experience, as much as you know I have an understanding of your experience. Intersubjectivity, a term whose meaning refers to this understanding between people and the process in which this understanding is formed, is of interest in this thesis.

This thesis asks in what way our similar embodiment is significant to how we *constitute* intersubjective understanding of the experience of the Other? In turn, how do we use our bodies to demonstrate this intersubjective empathetic understanding to one another?

Starting by considering intersubjectivity and empathy as a phenomenological concept this thesis explores the work of phenomenologists who have presented ideas about the manner in which intersubjectivity is implicated by embodiment.

1.3 Empathy

Empathy first appeared as a term relating to aesthetics, used to define the connection experienced when ones own feelings are projected into the dynamic structure of an object. Translated from the German Einfuhlung 'feeling into' by Edward Titchner in 1909. It was Theodor Lipps who first related empathy as extending to intersubjectivity between minds from a process of 'inner imitation', a grounds for understanding mental state or experiences of an Other. Empathy not only plays a part in our aesthetic appreciation of objects, it is understood as the primary basis for recognising each other as minded creatures rather than merely bodily forms and outward movements without psychological substance. In his projective theory of empathy, Lipps describes empathy as a modality of knowledge, as 'irreducible as our memory of past objects' and Lipps claims it cannot be reduced to inference or analogy to the Self as the mind

of the Other cannot be presupposed to be like exactly like our own mind. Rather empathy between minds is reliant on the inward projection of the gestures and expressions of interlocutors that indicate mental states and intentions, for Lipps it is this projective process that allows for interpersonal understanding. Lipps argued that empathy relies on projection because we only know our own experience, and thus stated that this explained the instinctual drive toward imitation and expression (Stueber, 2008).

Similarly for seminal phenomenologist Husserl the mind of the Other must be mediated through bodily presence- the inner mind of the Other is never directly perceived like the Self through inner consciousness (Husserl, 2006).

Max Scheler (1954) pushed this theory further by arguing that to observe the outward bodily expression of an experience is to perceive the essence of the experience itself in *expressive unity*, as opposed to body movements identified as representative of the experience. In this way we are empathetically able to experience the mind of the Other but as an *embodied* mind; like our own embodied experience. Thus empathy relies on expressive unity to form an embodied basis of our experience of the Other within the realm of psycho-physical phenomena (Zahavi, 2001; Scheler, 1954).

Schutz adds a layer of complexity to this idea of expressive unity; suggesting empathy can't be accomplished from an interpretative perspective constituted by different cultural presuppositions. Expressive behaviour is socially and culturally embedded so the mind of the individual must be constituted within this socio-cultural framework (Schutz, 1990).

1.4 Empathy between moving bodies

This brief consideration of phenomenological approaches raised the following questions: How do we use our bodies to communicate about embodied experiences that cannot be directly experienced by the Other, such as physical pain and affective states? How do we use our bodies to form an intersubjective understanding about these inner experiences of another?

Here we treat these questions as empirical, what do we physically do with our bodies to communicate inner experiences with each other? Although intersubjectivity requires an understanding between both communicating parties, this thesis approaches these questions by focussing on

the ways in which a *recipient* demonstrates their understanding through nonverbal interaction. Embodied or nonverbal interaction includes the use of expressive gestures and postures performed by interlocutors within an interaction.

1.5 Mimicry and intersubjective understanding

One general hypothesis is that mimicry of one another's expressive behaviour during interaction helps us form an intersubjective understanding. For example, to better understand how another feels in a situation they are describing to me, I mimic their expression to internally simulate and demonstrate my understanding of their experience with my body.

Mimicry broadly refers to the act of representing something (or someone) by adopting or reproducing it's (or their) characteristics. This thesis focusses on behavioural mimicry during nonverbal interaction. This is the morphological and dynamic matching of gestures, postures or expressions previously performed by an interactional partner. For example; tapping a foot after observing someone tap their foot. Behavioural mimicry is said to be spontaneous, nonconscious and automatic responses to expressive stimuli, based on a perception-behaviour link (Chartrand and Bargh, 1999; Preston and de Waal, 2002)- this theory postulates that the perception of a specific nonverbal behaviour automatically entrains the same behaviour in the perceiver. Behavioural mimicry is generally measured by frequency of imitated behaviours, rather than accuracy or temporal distance between the imitation and the imitated behaviour (Chartrand and van Baaren, 2009; Hess and Fischer, 2013). Behavioural mimicry is frequently correlated with affiliation and rapport, as it is said to synchronise interactants behaviour to foster smoother and more coordinated interactions leading to a firmer understanding of one another. Previous studies that search for empirical evidence of automatic behavioural mimicry are reviewed in Section 3.2.1.

Studies on facial mimicry go beyond stating that mimicry serves to synchronise behaviour during interaction leading to prosocial effects; Hatfield et al. (1994) suggest that the mechanism leading to the internal simulation of another's affective states is through the automatic adoption of associated facial movements, vocalisations, postures or bodily expressions. To mimic a sad face, for example, triggers one's own associated feelings of being sad; this is claimed to be the basis of emotional contagion (see Section 2.3.1 (Bourgeois and Hess, 2008; Sonnby-Borgström,

2003). This idea is consistent with research detailing a mirror neuron system; this identified a specific set of neurons that fire when an action is performed and when the same action is perceived (Iacoboni, 2009; Gallese, 2013).

By looking at bodily behaviour in conversation, this thesis analyses how well this simulative model characterises what people do when talking about one another's felt experiences. Investigating spontaneous interactions of this nature set the corpus apart from previous studies of behavioural mimicry- as they are specifically designed to promote an interactional situation where people share descriptions of common embodied experiences in order to encourage empathy. To test the hypothesis that behavioural mimicry has it's basis in simulative mechanisms promoting empathy; this work examines the similarities in gesture production and posture congruence in real conversations.

1.6 Empathy in actions

Our embodied resources can also be used in a more strategic manner. Empathetic feedback such as *Motor Mimicry*, the performative display of an expected expressive behaviour in reaction to an experience in the perspective of another, can be used to selectively demonstrate an understanding of the experience of another (Bavelas et al., 1986, 1987, 1988). These sorts of displays are not a straight mirror of the perceived actions of a conversational partner, but display how the other *would* behave in the situation that is being described. In other words, performing an expressive display from the perspective of the storyteller. For example, if you tell me you've hit your head and it hurts, I grimace and say 'ouch!' These displays are interesting as they not only show an understanding of the situation being described but also what *type* of experience it is for the other, or rather an understanding of how that experience affected the conversational partner describing it. This points to the importance of exploring how feedback is actually used in conversation. What types of feedback responses are simulative, such as relating the described experience to one that they have had themselves and how do they function intersubjectively. Here, motor mimicry is not regarded as a typical case of behavioural mimicry, as these feedback displays are not always performed after observation of the same behaviours. It is not a case of 'I wince after observing you winch' but 'I winch to demonstrate I understand an appropriate response to your experience in your perspective'. For this reason, the underlying simulative

mechanisms set out above are not attributed to these responses.

Other forms of embodied feedback, such as nodding or laughing, are thought to facilitate the incremental process of a conversation as a joint activity (Clark and Brennen, 1991; Bavelas and Chovil, 2000). Generic forms of feedback, like nodding, can first serve the function of showing attention and awareness but also declare comprehension. Specific forms of feedback, such as motor mimicry demonstrate a further level of understanding that address attitudinal or emotional responses. This study will explore the forms of feedback responses to the expression of embodied experiences by recipients and explore their implication for later proposed explanations of intersubjectivity.

1.7 Exploiting our embodied resources

An additional hypothesis developed here proposes that our embodiment can be used to provide a strategic auxiliary resource when there is a problem in conversation. This hypothesis can be addressed by analysing what happens when there is a difficulty. This can be readily identified by the occurrence of clarification questions or dis-fluencies leading to self-repair. How the interlocutors respond in this situation, in terms of feedback and mimesis, can reveal how embodied interaction is used to deal with misunderstanding.

1.8 Document structure

This document begins by reviewing literature on the concept of intersubjectivity in phenomenology, defining what an intersubjective understanding means when we encounter the Other in chapter two. This chapter addresses the questions: How is knowledge of another living consciousness constituted? Specifically, how is embodiment significant or implicated in intersubjectivity? In order to gain an insight on these questions, Husserl's intersubjective experience of empathy through pairing is explored. As a comparison Schelers' concept of expressive unity and Merleau-Pontys' concept of intercorporeality are also covered, arguing that intercorporeality, similar to Husserl's pairing, is defined by an intersubjectivity of bodily experience. For an alternative view Schutz' theory of 'reciprocity of perspectives' is explored, in particular the claim that intersubjectivity of bodily experience is mediated by a knowledge of a social world.

The concept of empathy and emotional contagion follows in chapter three. They are an important consideration for this study as they concern feeling another persons experience as a way of understanding it, *intersubjective* empathy as the knowledge of a mutual understanding of the said experience. The role of embodiment behind these experiences is highlighted in the review the how these concepts are empirically investigated in current studies looking at mimicry and interpersonal coordination in embodied interaction in both infant-carer developmental interactions and adult interactions.

Chapter four details the collection of two corpora designed to test these accounts. The first corpus of dyadic interactions contains conversations about embodied experiences, such as painful or pleasant bodily experiences like a headache or laughing. The second corpus of dyadic interactions contains conversations about current affairs, such as views and opinions of reality TV programmes and the impact of the Olympic games on the host city. When asking the question of how we use our embodiment to understand the experience of another, one way to provide an answer is in terms of the overt behaviour of the recipient. Focussing on the hypotheses highlighted above, what different forms of overt mimesis are produced in response to the embodied expressions of the felt experience being conveyed.

Chapters five, six and seven detail analyses that uses hand coded gestures, feedback and clarification questions, body movement data and a new approach to quantifying posture congruence. The analysis shows the embodied responses observed within these interactions are intentionally placed and formulated to facilitate the incremental process of a conversation as a joint activity. This is inconsistent with accounts that propose there is an automatic and non-conscious propensity for people to mimic each other in social interactions.

Chapter eight discusses six examples of excerpts that appear to contain behavioural mimicry, analysing how mimicked behaviour is organised within the structure of a conversation and drawing out the strategies people use to exploit a common embodiment. The concluding chapter addresses the questions raised in this thesis, outlining how the evidence fits together and provides a critical assessment of the thesis.

Part I

Literature review

Chapter 2

Concepts of intersubjectivity

The problem of intersubjectivity has been most consistently addressed by phenomenologists and the cognitive sciences. For both, characterisation of the relationship between Self and Other is a foundational problem: what is intersubjectivity and how is it possible to account for it in either phenomenological or cognitive terms? In both cases, embodiment has often been accorded a critical explanatory role.

2.1 Cognitive theories of mind

Theories of mind in the cognitive sciences are traditionally based on a Cartesian conception of mind, making the essential distinction between one's own mind, to which we are assumed to have direct, infallible access and the minds of others, where our access is indirect, fallible and (for some theories) inferential. How we constitute our understanding of the other as another minded being is the question behind the Theory of Mind debate. Two main camps have emerged in the attempt to explain this understanding, a Theory theory of mind and a Simulation theory of mind. Below gives a brief explanation of these different stances on a theory of mind.

2.1.1 Theory theory

'Theory of mind' theory (TT), put briefly, is the attribution of mental states to other people as a theoretical inference. TT claims we explain and predict another person's behaviour by relying on an innate or acquired theory of how people generally behave, and mental states such

as beliefs or desires that cause their behaviour. This theory is useful as it makes the distinction that other minds are not mere duplicates of our own, allowing others to be attributed a varying set of beliefs and intentions, of which can be different from one's own and each other (Morton, 1980; Gopnik, 2003).

Zahavi (2010) argues for the rejection of TT, suggesting that a knowledge of the *mindedness* of the Other cannot be based on analogy or inference based on theory because of three main issues; firstly, analogy is too complex to explain an infant recognition of other minds because of their lack of experience to formulate such a theory; secondly, similarity in the perception of both Self and Other is not close enough in experience, because of the lack of all sense modalities, such as proprioception; thirdly, TT doesn't explain empathy or emotional contagion, because these experiences are based on *feeling* the affective state rather than understanding the cognitive processes that lead to them (Zahavi, 2010).

2.1.2 Simulation theory

Simulation theory attempts to describe the skills used for ascribing, explaining and coordinating with others in the social world. Otherwise known as *mind-reading* and *folk psychology*, this theory suggests a mental representation or simulation of the mental state or mental processes of another person is generated in order to understand the mind of another. These resources are often termed psychological competence. The theory suggests the expressions of another are internalised or internally simulated to produce a vicarious experience of the mind of another. Simulation theory or folk psychology is based on the idea other minds are like our own (or rather ego-centric inference). One uses their own mind as a model to simulate the mind of the Other. The simulation extends to temporal and spatial relationships as well as emotional and affective states. Simulation theory claims we have no need for a *theory* because we have an inner model we can use for simulating another person's mental states, and this model is our own mind. Thus, we model the beliefs and intentions of others whom we deal with as if we were in their situation, or as if we were them. Suggesting this empathetic understanding does not need mental concepts or theories to work as it relies on non-conceptual simulative representation (Goldman, 1989; Gordon, 1986).

Goldman (2006) develops the theory further by suggesting there are two levels of simulation.

Firstly, low-level non-conceptual representations, the application of which leads to a vicarious affective or emotional mental state. This is internally simulating affective states, such as sadness or excitement, without an understanding of the circumstances that cause them. Secondly, high-level simulations referring to cognitive ascriptions such as role-taking or perspective-taking, this simulates an understanding of the effect that situational contexts have on the attitudes, intentions and affective states of another (Goldman, 2006).

As simulation theory is regarded as an empirical theory, it has been questioned whether it can answer the philosophical problem of Other minds. As the theory relies on our minds being the same as each other in order to simulate, it doesn't account for the different beliefs and intentions held by different people. Even simulating the beliefs and intentions of another using one's own mind as a model would not simulate the mind of another, but only one's own mind in that perspective (Gallagher, 2001; Gallagher and Zahavi, 2008; Gallagher, 2012).

Embodiment plays no essential role in these Cartesian models of intersubjectivity, often referred to as Cartesian Dualism, as these concepts separate the mind from the body. However, an important point to consider is while direct access is possible to one's own mind, access to the mind of the Other is mediated through their behaviour. Zahavi (2010) states intersubjectivity (or theory of mind) is a simultaneous analysis of the relationship between subjectivity and external reality. Individuals reciprocally illuminate one another conceptually, and can only be understood in their interconnection through interaction (Zahavi, 2010). What follows is a brief overview of key phenomenological ideas that analyse how we use our bodies to form an intersubjective understanding.

2.2 Phenomenological accounts of intersubjectivity

2.2.1 Husserl's notion of intersubjectivity

According to Husserl, intersubjective experience plays a fundamental role in our constitution of both ourselves as objectively existing subjects, other experiencing objects, and the objective spatio-temporal world. Husserl describes the nature of the Self by proposing the 'I' is not an experiential phenomenon by the thing experiencing, suggesting the Self is not the performance of behaviours or acts, or even character-traits but the Self is the thing possessing the characters-

traits as a property and is the thing that *performs* the acts and behaviours (Husserl, 2006).

Soren Overgaard (2006) interprets Husserl's notion of the Other as a minded entity that is not an inference nor a mere duplication of the Self, but the experience of the Other is mediated by a bodily representation of the Other. In other words the Other is only experienced through perceiving an outward bodily manifestation rather than through direct access to their experience (see Section 1.3). Husserl names this indirect experience a *presentiation*- the transcendental constitution of the Other. For Husserl, a universal co-existence leads to a subjectivity, both for myself and to others. The Self is constituted in an intersubjective reality, mutually shared in co-existence with each other. Subjectivity, cannot be separated from the experiential, no-one can possibly escape from their own mindedness or subjectivity and this is always experienced as a part of a 'community of subjectivities'. Husserl goes as far to say our bodily existence is something that makes intersubjectivity possible, as a mind embodied, performing acts 'indivisibly mental and bodily' and certainly not something established in and through inferences (Overgaard, 2006).

Husserl terms this intersubjective knowledge of co-existence *pairing*, the I-Other relation. To constitute a pair there must be two mutually contrasting variations of 'embodiment per se'. The Other is not a model or a replica of ourselves, but a pair in that causal relations are the same for one another, distinguished by an indirect experience of their effects. 'A 'here' that is transcendent of my own precisely because I experience it as 'there' in paired contrast of my own 'here'.' (Husserl, 1969, p. 115).

Notably for Husserl, the concept of intersubjectivity was not concerned so much with a particular understanding of others but with the transcendental conditions making this understanding possible. Husserl attempted to resolve how we constitute the Other's body as a 'lived body' by proposing Intersubjective experience is empathetic experience. Husserl adopted the term empathy from Theodor Lipps and the Munich school, but in a modified form. Empathy occurs in the process of our attribution of intentional acts to other subjects and forms the basis of an experience of Otherness (Husserl, 2006).

What is interesting is how Husserl treats empathy as constitutive of intersubjectivity but also as the precondition for knowledge of an objective reality or external world. Through experiencing an Other's experience of the world, one sees the world outside subjectivity, one can then

experience that world as an objective one, one larger than a singular experience of it. Husserl refers to *foreign subjectivity*, as regarding the Other as a minded being, with a mind separate from our own and the bearer of different perspectives, and it is this that allows one to constitute an objective reality (Husserl, 2006). The intersubjective process Husserl describes is dependent on the reciprocal transformation of an embodied relationship;

> In order to establish a mutual relationship between myself and an other, in order to communicate something to him, a bodily relation must be instituted, in empathy I participate in the other's positing (Husserl, 2006, p. 25).

The major criticism of Husserl's notion of intersubjectivity is its transcendental approach, as will be discussed in the following section covering Schutz' theory of 'reciprocity of perspectives.' Husserl regarded the mind of the Other as something that could never be known directly. The Other's experience can never be shared in the same direct manner as the Self is. Husserl was bothered by this sense because he believed it was only in knowing the Other 'absolutely', that disproved the notion of solipsism. Husserl felt that Cartesian Dualism obscured the point and that the 'nature of experience is not strictly mental or emotional but transcendental, in that it offers the possibility of recognising the world, not as a construct, but as a phenomenon' (Overgaard, 2006).

2.2.2 Expressive unity and intercorporeality

Dan Zahavi (2010) synthesises some relevant points with regard to intersubjectivity. Starting with Husserl's reliance on embodied experience in the process of empathy, the bodily exploration permitting an experience of one's own exteriority and the constitution of intersubjective understanding (in a theory of mind sense). Zahavi adapts this notion of intersubjectivity by turning to Max Scheler, Scheler rejects the assumption that the point of departure is one's own self-consciousness and the mind of the Other is inaccessible. Scheler states that to see another's outward emotion in their overt behaviour is to *perceive* them implicitly, as a unified entity. *Ausdrucksbewegung* or expressive movement is a bodily movement or sound that serves as an expression of a person's emotional state of mind. Zahavi (2008) regards bodily gestures and behaviour as expressive phenomena, and not simply as physical movements. The experience of another is not inferred but directly experienced, that intersubjectivity is not the process of

perceiving meaningless behaviour and attributing intellectual psychological meaning, but the Other is an 'embodied mind with expressive unity'.

> Affective and emotional states are not simply qualities of subjective experience, rather they are given in expressive phenomena and this presents us with a direct and non-inferential access to the experiential life of others the primary datum and primitive stratum (Zahavi, 2008).

Zahavi suggests that the relation between the other mind and body is more meaningful than our perception of inanimate objects and that the otherness of the other is exactly manifest in their elusiveness and inaccessibility (Zahavi, 2010).

Merleau-Ponty (1996) regards the body as important in the constitution of the Other, under the caveat that there is a distinction between perceiving the body as an object, or the natural puppet body: *korper* to perceiving a living entity: *leiber*. He states that this distinction relies on experiencing a cultural world of one's own, one must experience being among beings, to understand the significance and intention of the perceived gestures of the Other. Whose manifest behaviours are interpreted by analogy to one's own inner experience; 'If I am inhabited with a consciousness, then others can be as well' (Merleau-Ponty, 1996).

Importantly the *we*, as a collectivity of consciousnesses, is not a plurality of the *I* but as separate psycho-physical objects of the same type. Merleau-Ponty states that 'My body and the Other are two sides of the same phenomenon' (Merleau-Ponty, 1996), suggesting that Others are inhabited bodies with a body schema[1] *like* one's own, but not the same as one's own.

> The reason why I have evidence of the other man's being-there when I shake his hand is that his hand is substituted for my left hand, and my body annexes the body of another person in that 'sort of reflection' it is paradoxically the seat of. My two hands 'coexist' or are 'co-present' because they are one single body's hands. The other person appears through an extension of that co-presence; he and I are like organs of one single intercorporeality (Merleau-Ponty, 1996).

Merleau-Pontys' term *incorporeality* refers to this mutual understanding of the body as bearer of expressions and intentions reciprocal to ourselves. This intercorporeal relation within an

[1]Body schema is a technical term referring to how one understands their body in relation to the external world

embodied interaction is primary, as it requires no intellectual interpretation like theoretical inferences or simulations but, as Scheler also points out, is a direct relation between people. For example, when one perceives an expressive gesture of pain, it does not merely signify the pain it embodies but *is* the pain.

Gallagher (2001) argues that intersubjectivity comes out of 'a common bodily intentionality that is shared across the perceiving subject and the perceived other.' So the other mind is revealed to us through the embodied action that expresses the intentions of the other. Gallagher (2012) further argues that expressive unity alone is not enough for intersubjectivity, that to understand the Other is to apply their behaviour to their narrative, their situational and motivational context in order to intersubjectively make sense of it.

Extending on this, De Jaegher (2009) argue that social cognition emerges from embodied social interaction. It is suggested that social cognition does not involve autonomously interpreting another's actions but is manifested in a dynamic process of interacting and coordinating with others.

2.3 Concepts of empathy and emotional contagion

Although the above sets out the philosophical nature of how intersubjectivity and empathy can be constituted a consideration of the practicalities of this knowledge of another's experience follows with a discussion of the concepts of emotional contagion and empathy.

Emotional contagion occurs when people start feeling similar emotions caused by passive exposure to other people. The concept of emotional contagion does not require awareness that the transmitted affect has its source in others, they may be experiences without any recognition of sharing involved. Empathy by contrast requires knowledge that the transmitted affect has it's origin in another. This knowledge leads to an understanding of the experience of another through feeling how they feel.

Anne Jacobson (2006) suggests that even if the assessment of the Other's mental states are not accurate, one may still find emotional contagion influenced by the beliefs and other's attitudes without understanding what the shared emotions actually are. Furthermore, this lack of contextual stimulus suggests that the emotional representation provokes emotional arousal

before contextual engagement (Jacobson, 2007). It could be suggested that empathy is actually an extension of emotional contagion plus the knowledge that this transmission has come from another.

2.3.1 Emotional contagion: Embodying expressions

How does transmission occur? Hatfield et al. (1994) suggest that the mechanism leading to the induction of emotions is through the automatic adoption of associated facial movements, vocalisations, postures or bodily expressions. To mimic a sad face, for example, triggers one's own associated feelings of being sad. Their studies show that when watching a stimulus video of an actor describing a sad or happy experience, subjects tend to mimic the facial expression the actor, and self-reports of their own emotions after having mimicked the facial expression show a correlation in emotion with the either the sad or happy experience. Primitive emotional contagion is described as 'the tendency to automatically mimic and synchronise expressions, vocalisations, postures, and movements with those of another person's and consequently, to converge emotionally' (Hatfield et al., 1994).

It is proposed that the grounding of knowledge is a matter of reliving the past experiences that were encoded at the time of acquisition, including sensory, motor and affective modalities. So when it comes to emotional information, the bodily expression of that emotion is encoded in its neural representation forming a reciprocal relationship between overt behaviour and emotional states (Hatfield et al., 1994).

Sonnby-Borgström (2003) tests the notion that automatic physiological mimicry leads to unconscious matching of affective states. She exposes participants to sad, neutral and happy facial expressions on a computer screen for varying short intervals to produce different degrees of cognitive processing. If the facial expressions shown on the screen were displayed for an interval of 17-40ms then only a pre-attentive awareness of the expression were assumed to be acquired. Participants reported how they felt after the exposure and their facial expressions are monitored via Electromyography (EMG)[2] measuring facial muscle activity. The automatic mimicry mechanism should mean that emotional contagion and facial mimicry occur even at a pre-attentive level, before conscious or focal awareness. She found supporting evidence for this

[2]Facial electromyography measures muscle activity by detecting and amplifying the tiny electrical impulses that are generated by muscle fibres when they contract.

in the facial muscle activation and self-reports of the participants at a pre-attentive level *and* with conscious awareness.

Hsee et al. (1992) ran a study to test whether emotional contagion was induced primarily from a mimicked bodily expression of emotion over verbal descriptions. The experiment showed subjects videos of interviewees describing happy and sad experiences but with an inconsistent computer voice-over to facial expression. Although subjects rated the emotion in verbal description to be what the interviewee must be feeling, the self-report of how the subject was feeling had more correlation with the nonverbal behaviours exhibited in the stimulus videos, providing positive evidence for their theory of primitive emotional contagion. In the results of experiments on the influence of imitating facial expressions to decode the emotion behind them, Hess and Blairy (2001) observed that dynamic facial expressions do promote emotional contagion and mimicry, but no evidence to suggest that the mimicry was the *facilitator* of the emotional contagion and no evidence that observing or mimicking a facial expression is related to accurately decoding the corresponding emotion. It is suggested however, that mimicry provides the mimicker with an understanding of another's emotional state. They suggest that observers mimic emotional expression as a mechanism to decode the emotion behind an expression, but do not find evidence pointing to the idea that observers reflexively adopt the emotional displays they see because of behavioural imitation per se.

A further study shows facial mimicry contributes to a quicker recognition of affect of the facial expressions of others. Stel and van Knippenberg (2008) show when participants are instructed to keep their faces still, therefore exhibiting less facial mimicry, they are slower to identify whether the expressions on photographs are happy or sad, than when instructed to keep only their shoulders still. In line with this, a study by Sonnby-Borgström (2003) shows facial mimicry was more prevalent in participants testing higher for emotional empathy[3]. Facial mimicry has more recently been shown to improve a perception of emotions (Neal and Chartrand, 2011; Ponari et al., 2012; Sato et al., 2013), even correctly reading true and false smiles (Rychlowska et al., 2014; Krumhuber et al., 2014), however mimicry has negative consequences for detecting lying (Stel et al., 2009). A distinction between emotional contagion, empathy and emotion recognition is made here, recognising that a sad face indicates sadness in the person behind it is

[3]Emotional empathy was measured using the questionnaire measure of emotional emotions (QMEE) (Chlopan et al., 1985)

a different experience from feeling the sad emotion via emotional contagion. These studies seem to suggest that empathy can be experienced when an emotion is transferred by an automatic imitation of facial expressions and then associated with the bearer of the expression.

Niedenthal and Barsalou (2005) ran studies evaluating the associated emotion of certain words and videotaping each subject's facial expression when answering, their results showed a tendency to adopt the facial expression of associated emotion that was given to the word, for example, pulling a disgusted expression when evaluating the emotion associated with the word 'slug'. Further support was shown when they asked a second set of subjects to evaluate whether a word was in capital letters, these subjects did not embody the associated emotion of that word. To explain this, they use the notion of 'embodied architectures' suggesting that knowledge is grounded in bodily states and analogous to modality-specific systems of perception structures. These include the sensory systems that form perception, the motor systems that give rise to actions and the introspective systems underlying conscious experience of emotion, motivation and cognitive operations. Calling this the 'perceptual symbols system theory' (PSS theory): 'simulations of perceptual, motor, and introspective experience underlie the representation and processing of knowledge. So during knowledge use, the multiple modality-specific regions that encoded the experience are activated, rather than the 'final' abstract regions at the end of processing streams. They propose that: the 'understanding of another person's emotional state, comes from mentally 'recreating' this person's feelings in ourselves' and accordingly embodiment has an essential role for intersubjectivity (Niedenthal and Barsalou, 2005; Niedenthal, 2007).

Similar to Niedenthal's embodied architectures, Varela et al. (1991) explore the notion networks of different sensorimotor patterns (motor and neural programs) are coupled with meaning, suggesting concept and meaning arise from unique neurological configurations. For example, a certain configuration of neurons firing at once has a unique representation, and these are said to represent conceptualised understandings of perception. They also state it is in our experience of the world these configurations are structured, and this is why meaning is always mediated by embodiment, cultural embeddedness, imaginative understanding, and tradition of historical views on rationality (Varela et al., 1991).

Soto and Levenson (2009) describe the mechanism between mimicry and contagion as *physiological linkage*. mimicking an expression of emotion the patterns of physiological response

(autonomic and peripheral nervous system activity) in the observer mimic those of the person being observed, leading to emotional contagion. They argue for a similar emotion to be 'caught', a similar expression and physiological activity should be associated. Support for this could be found if people would more accurately decode the expressions on same-race faces than other-race faces - and differences in the expression of emotions between cultural groups could lead to less accuracy across cultures. Their experiment asked subjects from four different cultural backgrounds to rate the emotion of an interaction between a dating couple in four different stimulus videos, where the couple in each video were from the four different cultural backgrounds. However, their findings found no decrease in accuracy if decoding emotions from different cultural groups (Soto and Levenson, 2009).

In summary, the mechanism underlying emotional contagion is claimed to be an automatic propensity to mimic one another's expressive displays. This has been largely researched for facial mimicry; mimicking a sad facial expression is claimed to trigger the associated feelings of being sad as the feeling state is neurally encoded with overlapping sensory, motor and affective representations. This forms a reciprocal relationship between the expressive display (a sad facial expression) and emotional states (feeling sad) (Hatfield et al., 1994). Sonnby-Borgström (2003) find evidence that imitation of facial expressions automatically happens at a pre-attentive awareness; for example, imitation occurring after exposure to a photograph of an expression for such a short interval the perceiver is not consciously aware they observed it. Moreover, Hsee et al. (1992) provides evidence that the affective state triggered via facial mimicry of expressions in stimulus videos has more influence over the affective state expressed in an inconsistent voice-over. Facial mimicry has been found to improve and increase the speed of perceiving emotion in others, to the point that it impedes the detection of lying (Stel and van Knippenberg, 2008; Sonnby-Borgström, 2003; Neal and Chartrand, 2011; Ponari et al., 2012; Sato et al., 2013; Stel et al., 2009; Rychlowska et al., 2014; Krumhuber et al., 2014). Niedenthal and Barsalou (2005); Niedenthal (2007) notion of 'embodied architectures' suggests that concepts and meaning are rooted in bodily experiences, a similar expression and physiological activity are associated to form unique neurological configurations related to the experience. This is described as 'networks of different sensorimotor patterns (motor and neural programs) by Varela et al. (1991) and 'physiological linkage' by Soto and Levenson (2009).

2.3.2 Empathy: Two-step models

The Empathy-Imitation Russian Doll model based on de Waals work with primates describes this process as automatic and involuntary mirroring of the mental activities or experiences of another based on the observation of their bodily activities or facial expressions, ultimately implying the origins of empathy lies in an innate disposition for mimicry. Which describes 'an inner tendency giving rise to similar kinaesthetic sensations in the observer as felt by the observed target' (Preston and de Waal, 2002).

de Waal (2010) describes his notion of empathy as the capacity to:

- be affected by and share the emotional state of another
- assess the reasons for the other's state
- identify with the other, adopting his or her perspective.

de Waal (2010) notes empathy has two sides: One is higher-level, such as adopting the perspective of another which is emphasised in psychology and leads to the view empathy is a higher-level cognitive system. The second is a more primitive, pre-linguistic form of empathy displayed by infants and some primates.

> I consider this emotional side the core around which all empathy is constructed. Empathy is first of all a bodily and emotional connection. If this core is lacking, you can try to imagine another's situation as much as you want, you would lack investment in it. You might take the other's perspective, but we wouldn't call it empathy if you stay emotionally aloof (de Waal, 2010, p. 10).

Enz et al. (2009) proposes this two-type (affective/cognitive) empathy distinction is more akin to a continuum. At one end of the continuum sits the affective non-cognitive processes and at the other end sits advanced cognitive processes, with simple cognitive processes in the centre (Enz et al., 2009).

Karsten R. Stueber (2006) presents a non-simulationist conception of empathy in his work *Rediscovering Empathy*. Describing another double-tiered approach, firstly a low-level basic empathy based on inner imitation and a high-level reenactive empathy based on cognitive imitation. Agreeing with de Waal 'our relation to others is not primarily a theoretical matter, but

rather something rooted in our emotional life' (Stueber, 2006, p. 87). However, Stueber suggests a perceptual theory of other minds, rejecting de Waals proposal that emotional contagion or mimicry are the basis of empathy but also disagrees with the suggestion that empathy is based on a direct association between cognitive interpretation of the Other and past experiences (Stueber, 2006).

2.4 Schutz and practical intersubjectivity

This section considers the implications of Husserl's general approach for the observable use of embodiment in interaction; not as mimicry and simulation but as a potential resource for 'defending intersubjectivity'.

Schutz (1990) rejected Husserl's transcendental approach to intersubjectivity. Husserl's notion of empathy occurred through the process by which one transferred an understanding of the Other as a 'lived body' to another. The verification the Other's living body was like one's own could be achieved if it manifested outward behaviour congruent with what one would expect of a minded body, that it is 'like' our own. Schutz argued instead the experience of the Other's bodily presence is always from the outside, and cannot correlate with the internal experience of the Self to sufficiently constitute a transfer, Schutz argued this verification could only be achieved through a correlation of the Other's behaviour with 'what was 'congruent' behaviour drawn from the social-world[4]; and presuppositions of how bodies ought to behave' (Psathas, 1973; Schutz, 1990).

The *social world* for Schutz, was the main foundation for Intersubjective understanding. Turning to the concept of a *lifeworld*, (which is identical to Husserl's notion), the lifeworld represents one's own perspective as mediated culturally, socially and evolutionarily. This is described by Schutz as:

> sets of assumptions, beliefs, and meanings against which the individual judges and interprets everyday experiences (Schutz, 1990).

Schutz suggests an understanding between individual lifeworlds is possible only because of the

[4]Schutz regards the social-world as constituting a 'world of experience from the outset as an intersubjective world of culture. As men among a world of men.' (Schutz, 1990)

enactment of 'roles'. In the words of Psathas: 'Society generates its own categories and with them, its own criteria of what is intelligible to one another by taking established 'roles' [5].' (Psathas, 1973, p. 65). Schutz proposes three principles for understanding behaviour:

- if people are to find behaviour interpretable then it must fit a 'standardised story'.
- interpretability is something agents always seek.
- 'novel' behaviour is a part of social change.

Schutz' theory does not rely solely on in the overt behaviour of the Others to constitute intersubjectivity, but in the projection of actions of the Other. Schutz operates his theory on the premise the Other's intention can only be known if one can make the assumption there is a common scheme of reference in how we express our intentions in overt behaviour. Schutz proposes 'When I seek to understand another's behaviour in an ideal-typical fashion, a twofold method is available to me. I can begin with the finished act, then determine the type of action producing it, and finally settle upon the type of person who must have acted in this way. Or I can reverse the process and, knowing the personal ideal type, deduce the corresponding act.' (Schutz, 1990, p. 25). In short, the theory is the belief of the Other is understood by rules of interpretation, explanation and justification specific, or the shared intentions, particular to a society.

We can now turn to Schutz' Reciprocity of Perspectives: The mutual assumption the objective world, including objects and events, can be experienced by ourselves and to others respectively. Such things are similar in the way they are empirically presented, but whose *meaning* is something different since I take a reciprocal perspective to another. For example, I am 'here' (physically, socially, morally and so on) and the Other is 'there'. Schutz considers the interpretation of expressive behaviours to be from a *vernacular of collectivity*, so in Schutz' sense intersubjective agreement and consensus is not a condition but an ideal standard of conduct. This is how we interpret and understand the bodily behaviour of the Other.

In the domain of conversational analysis, the notion of intersubjectivity is of importance to social order. Schegloffs (1992) approach takes intersubjectivity as forming a 'mutual understanding' between interacting parties and not the 'problem of other minds' as discussed above. Schegloff argues that communicating is a complex process, with many different levels, 'remem-

[5]Schutz regards 'roles' as one's biographically determined situation.

bering what has been said: deciding what to say next in light of overall conversational goals; correcting misapprehensions, and reinterpreting past interactions all contributes to the complexity of this process. Like Schutz, Schegloff is arguing individuals must conform to 'common culture' or a sense of 'social norms' both in the process and performance of the interaction itself. Schegloffs approach originates in Garfinkels early ethnomethodological work. which introduces this concept of *social norms*. Garfinkel (1964) postulates *social norms* are regulatory not just for the purpose of 'social order', but the process of interaction or the interpretation of actions to be mutually understandable at all (Garfinkel, 1964). These norms can be related to Schutz' ideal standards of conduct, the presumption which matters when people explain deviations from it. Schegloff sums this up with the statement, 'without the world known in common one has not a misunderstood world, but no conjoint reality at all' (Schegloff, 1992).

However, if the common 'norms' or 'propositions' held at either end of the interaction party are not composed of the same substantive properties, then misunderstanding can occur - an example Schegloff offers is there could be many meanings for the same string of words, and a misunderstanding occurs if there is a difference in interpretation. Schegloff claims it is only in interactional, sequential coordination by which such misunderstandings are resolved and joint understanding or intersubjectivity is defended. Schegloff's analysis of their notion of intersubjectivity suggests intersubjectivity is more than the alignment of beliefs, knowledge or interpretations of the world, but 'a convergence between the 'doers' of an action or bit of conduct and its recipients, as co-producers of an increment of interactional and social reality' (Schegloff, 1992).

Although Schegloff's discussion is primarily based on the dialogue; the use of other local resources such as gesture are also considered to hold the potential to defend intersubjectivity in the same way.

2.5 Summary

In summary, Husserl's concept of 'pairing' suggests that by being similarly embodied we have the capacity to understand each other's experience through our manifest behaviours. Through 'pairing' an understanding of the experience behind the behaviour is available without the need to analogise or inwardly relive one another's experience through simulation. Note that this

concept of 'pairing' doesn't assume that we are identical, just that we possess a mind that is similarly embodied, subject to that same causal relations and possessing a stream of perception like ourselves. An intersubjective understanding is key here, only by recognising that others understand us in the same way can 'pairing' be fully constituted. In this sense, communication is crucial to intersubjectivity and empathy, an understanding of embodiment is used to reciprocally illuminate similarities during interaction will explain how this intersubjective understanding can be reached. So, although 'pairing' does not provide an infallible understanding of one another's expressions- it allows the possibility to make an interpretation that expressions are of a certain type.

Here, Scheler's notion of expressive unity, that an embodied expression and its associated affective state are not perceived as separate but perceived as essences of the same phenomena, extends on how embodiment is intregral to Husserl's concept of 'pairing'- previously Husserl lays the groundwork that specifies how a capacity for intersubjectivity can be possible through a recognition of Other's as 'objects of the perception'. Expressive unity is the notion that manifest expressive behaviours are understood in direct relation without requiring an anologous or simulative interpretation (Zahavi, 2008, 2010). For Merleau-Ponty (1996), co-presence (or intercorporeality) is an important aspect for intersubjectivity, stating that the reflection of embodiment perceived in Others reveals that another holds the capacity to experience as we do. Gallagher (2012), in line with Schutz (see below), postulates that this capacity is not enough for intersubjectivity, the perception of another's expressive phenomena is understood as as bounded by a contextual narrative, which to De Jaegher (2009) is only possible via the dynamic and joint process of interacting.

Studies of facial mimicry that dominate research on emotional contagion all take their evidence from participants responses to photographs or video as stimulus. As participants in these studies passively observe expressions made in photographs or videos it does not indicate whether or not facial mimicry is used when people are actively engaged in interaction. Embodied interaction, according to the key concepts highlighted by Phenomenologists Husserl, Merleau-Ponty, Scheler and cognitive theorists Gallagher and De Jaegher, is the primary the component that makes intersubjectivity possible, so these particular facial mimicry studies only present one side of the process. Furthermore, there are studies that show some expressions show counter-mimicry responses, such as a neutral expression in response to anger (see Section 3.2.4.1) (Iacobini

et al., 2009, 2010; Bourgeois and Hess, 2008). The two-step models of empathy discussed below attempt to bridge these gaps by considering how empathy can arise from both expressive displays and contextual and situational information.

Two-step models of empathy suggests that the expressive phenomena manifest of Other's is interpreted by two distinct processes; firstly, a higher level cognitive system that considers the perspective of the other in terms of situation and context, secondly, a more primitive, pre-linguistic consideration that arises out of an inner imitation of manifest behaviours (de Waal, 2010). The lower level could be argued to correspond to the mechanism underlying emotional contagion as discussed above; that an innate disposition to mimic one another's expressive (facial) displays triggers the associated feelings of behind the expression as the sensory, motor and affective neural representaions overlap. However, Enz et al. (2009) suggests that these processes are more akin to a continuum, and furthermore Stueber (2006) two-step model incorporates the idea that it is still possible to understand the affective dimension of another's experiece wholly based on higher, more cognitive, processes. These models highlight that empathy is considered to be an understanding informed by different sources available through interaction with Others. Ranging from an inner imitation of another's affective dimension to cognitively placing oneself in another's situation.

Intersubjectivity requires an additional component for Schutz, who places an importance on a shared social world on understanding one another's experiences, suggesting that each member of a social collective adhers to a role that allows other's to form expectations of their motivations or intentions that influence a conception of empathy. This is in line with Gallagher (2001) in that intersubjectivity is considered beyond the expressive qualities found in overt behaviour but is understood in context with a narrative that sits on a 'vernacular of collectivity'. Schegloff (1992) applies this social conformity to reside within a common culture that is essential in the process of interaction. In other words, a shared vernacular can only be possible with the existence of what Garfinkel (1964) terms social norms; he observes that a mismatch in these systems is cause for misunderstandings between interlocutors and uncertainties in other's motivations. Out of this, what is important for understanding between interlocutors is the sequential coordination of interaction that defend intersubjectivity via a shared structural communicative system.

2.6 Conclusion

Embodiment is considered as important for the attempt to conceptually resolve the problem of Other minds by phenomenologists. Ideas such as Husserl's notion of 'pairing' and Merleau-Ponty's definition of 'intercorporeality' present an alternative to the 'theories of mind' discussion in the cognitive sciences. Highlighted is that the capacity for intersubjectivity is possible through the reciprocal experience of another's embodied expressions during interaction- this mutual reflection of understanding illuminates one another as an 'object of perception' as we are. Much of the literature on emotional contagion and empathy postulate theories on imitation as the mechanisms that allow an inter-correlation between Other's expressions and the associated feeling state to vicariously understand Other's experiences. What is missing from these early facial recognition studies and two-step models of empathy is an analysis of how mimicry operates within the organisation of interaction; this would provide insight into how the recognition that others understand us in the same way we understand them.

Additionally, Garfinkel, Schegloff and Schutz see embodiment as a resource for helping to resolve the practical problems that intersubjectivity present, in the sense that embodiment presents local resources as a means of defending intersubjectivity. A further review of empirical studies that examine how embodied resources are used for intersubjectivity and empathy follows in Chapter 3.

Chapter 3

Empirical studies on intersubjectivity

The previous chapter discusses how the capacity to understand the experience of the Other is possible by experiencing expressive phenomena exhibited in others through embodied interaction. The discussed theories of emotional contagion and models of empathy attempt to explain the specific process in which an understanding of another's experience can be established rely on mimicry of expressive displays. This chapter explores empirical research on the role of embodiment in empathy and intersubjectivity, asking how we use our bodies in interaction to form and demonstrate our understanding of the experience of another; in particular how / if mimicry is part of this process. The first section analyses infant-carer nonverbal interactions; specifically weighing up the evidence relating to mimicry in neonates, given that a perception-behaviour link proposes that automatic mimicry is innate, it should operate from birth. Discussed is the theoretical basis of how (and if) mimicry can contribute to developing a mutual empathetic awareness between infant and carer. The section then turns to alternative explanations of what observed interactions between infant and carer could intend to express and how these exploratory interactional procedures contribute to infant's development of empathy and intersubjectivity.

The second section analyses adult interactions that detail evidence of behavioural mimicry, outlining the mechanisms that account for it; including theories from cognitive neuroscience and the social implications of behavioural mimicry in interaction. Nonverbal behaviours like

gesture and posture are examined in terms of how mimicry can fit into their function within interaction as well as feedback behaviours.

3.1 The development of intersubjectivity in infant-carer non-verbal interactions

Developmental psychology investigates the development of intersubjectivity within infant-carer interactions from the very first days as a neonate, up until 2 years, the point at which infants generally begin to verbalise more competently.

The importance of studying interactions at the pre-verbal stage is the opportunity to focus on nonverbal communication without having to consider the influence of concurrent verbal language as would be the case when studying adult interactions. The following literature places an importance on embodied expression between infant and carer to account for intersubjectivity, assuming that the affective states represented through expressive embodied behaviour is the first communicative contact. For this reason, nonverbal expression could be a fundamental vehicle for intersubjectivity not only at the developmental stage but as a foundation of the communication of affective states in adult interactions.

3.1.1 Evidence of innate behavioural mimicry

Meltzoff and Moore's (1977) studies on 12-21 day old infants found imitation of lip protrusion, tongue protrusion, mouth opening and finger movements of their carers (Meltzoff and Moore, 1977). They observed that infants were able to imitate postures and facial expressions in a manner that was cross-modal: where an expression an infant visually observes is translated to bodily imitation.

The authors reject the proposal that imitation observed is an innate releasing mechanism (IRM): a reflexive pre-programmed response or fixed action pattern, such as responding with tongue protrusion to nipple like stimulus. This is because the imitation is not stereotypical enough, the movement is never exactly the same and the range of behaviour can only be narrow and selective if the action is caused by an IRM. Nor does the responding action occur consistently enough to be a fully fledged reflexive action.

Later, Meltzoff and Brooks (2001) suggest that this experience of imitating others provides the infant with a 'discovery mechanism' for an embodied understanding of others. The authors indicate that this discovery mechanism must rely on three starting states:

1. A recognition of equivalences between perceived and executed acts
2. A recognition that bodily acts are systematically related to mental states
3. and 'like me' behaviour leads to the projection that Others also have mental experiences.

Infant imitation suggests an overt representation of the acts of others, a performed equivalence to the other and their own cross-modal tactile-kinaesthetic self. A line of argument that makes the recognition of self-other equivalences the starting point for intersubjectivity. This allows for the possibility that neonates are born with the ability to make the distinction between the Self and Other, through an innate 'theory of mind'. They argue that the reciprocal ability for imitation allows infants to learn to ascribe psychological properties to people by matching shared behavioural states. This suggests infants learn about others by a comparison to the self: 'like me', or for Husserl *pairing* as the foundation of intersubjectivity:

> Starting state nativism embraces innate imitation and cross modal mapping between self and other and suggests that infants can elaborate their understanding of others by processing them as 'like me' (Meltzoff and Brooks, 2001, p.175).

Meltzoff and Brooks (2001) later extend their explanation for a starting state nativism by suggesting that the process of imitation as provides an 'interpretative framework' to understand the behaviour of others by mapping their behaviours onto their own in order to understand the intention behind it. Further proposing that infants are born with the ability to recognise goal-directed acts, observing gaze following behaviour as evidence that infants understand when carers intend to look at something.

> Infants begin by interpreting human behaviour in terms of acts and goals, not muscle movement or physical motions, this seems to be the starting state (Meltzoff and Brooks, 2001, p.188).

This is supported with the observation of infants adapting their movements to reach inferred goals rather than simply imitating behaviour of their carers that would not function to achieve the goal. They observed after 9 months infants appear to have an understanding of even

unfulfilled goals attempted by carers and go on adapt their behaviour to achieve a goal that was not successfully acted out initially by the carer.

The main argument is that infant imitation can be regarded as the pre-verbal common code of human actions, that it is the only faculty the infant has to connect with others (Meltzoff, 2002).

Innate imitation is frequently supported by appeal to the *mirror neuron* system (see Section 3.2.3) that observed actions are understood in terms of one's own action programmes allowing infants the ability to imitate.

Ferrari et als. (2012) work observes the mirror neuron system in terms of a mu rhythm desynchronisation that is apparent from electromagnetic signal recorded from EEG scans in new-born macaques. An electroencephalogram (EEG) is a test that measures and records electrical activity of brain waves using electrodes placed on the scalp over various sensorimotor regions during normal neural activity. Mu rhythm is a type of emitted brain wave that can be measured via EEG. The mu rhythm band is proposed to indicate the conductance of synchronised activity in large neural groupings in the motor cortex (Pfurtscheller et al., 1997). The desynchronisation of neural activity, referred to as attenuation, is indicative of significant activation in that particular region in studies of humans and primates. Both the execution and observation of actions have been shown to result in an attenuation of the mu rhythm, reflecting an innate activity of the mirror neuron system in new-born macaques(Pineda, 2005). The mu rhythm desynchronises when a new-born macaque performs appropriate matching behaviour such as facial gestures (lip smacking) known to be communicative (Ferrari et al., 2012). Simpson et al. (2014) suggest that *if* this result is positively tested for in new-born infants it could be concluded that this mirror mechanism operates from birth; therefore born with the ability to engage in social interaction. In this sense neonatal imitation is a behavioural measure of the MNS, that contributes to perception-action equivalences that allow for these cross modal matches (Simpson et al., 2014).

Decety and Chaminade (2003) take a stance that draws from further work on the aforementioned MNS, drawing from studies on infant imitation they suggest that knowledge of the mind of the Other is based on an innate analogy that Others are like the Self. Decety and Chaminade look at evidence at a neural level, proposing the shared representations model. This asserts that the

actions and emotions of others are understood in terms of our own motor system programmes. For example, while observing a smile, the observer would mimic same facial muscles at a sub-threshold level and this would activate the same neural network in the observer and cause them to feel happiness. Here, it is suggested that the MNS is the neural mechanism that leads to emotional contagion as discussed in Section **??**

> Indeed, the vertebrate brain was designed primitively for governing motor activity with the basic function of transforming sensory patterns into patterns of motor coordination. Herein lies a fundamental basis for the interpretation, direct or indirect, of all higher brain processes including mental functions. Thus, in relation to perception the motor system plays an elementary role in social cognition, and intersubjectivity develops out of overt action (i.e., transactions between self and other) (Decety and Chaminade, 2003b).

This suggests that infants not only base their understanding of the Other on the imitated movements as above but incorporate goals and intention into their representation, through the process of interaction, intersubjective understanding emerge (Decety and Chaminade, 2003b).

3.1.2 Discussion of the evidence behind infant-carer behavioural mimicry

In a recent review Beebe et al. (2003) describe Meltzoff and colleagues explanations for the presence of neonate imitation as the neonate encodes events as a non-modality specific representational code. Non-modality specific simply refers to 'The perception as well as the production of human action are both represented within the same framework' - visual, spatial and temporal components of the expression that are cross modally available for the infant and the infant reproduces them as the first sense of 'You are like me' (Beebe et al., 2003).

One problem with this series of experiments is that the structure of the interaction was unnatural. When the infant was engaged (which was not all the time) they were observing faces repeatedly doing one type of facial expression. At the moment of the infant observing the expression they had a pacifier in their mouth, which was released at an interim period between stimulus stage. The facial expressions the infant made in the interim 'imitation' stage, were regarded as mimicry only if the expression matched the expression of the carer / experimenter in the previous stimulus stage. This entails an unusual, forced interaction that does not serve

any interactional purpose except for the purposes of the experiment. The infant cannot entrain temporally and any mimicry observed would not support the 'all modalities operate at once as a package' argument that the authors propose is the mechanism behind the phenomenon of cross modal matching. The idea that neonate imitation stands as being 'the origin of pre-symbolic intersubjectivity and an innate theory of mind' is not consistent with the evidence found in this study as the interaction is somewhat one sided. How long the infant has to interact is controlled, the feedback they receive because of their actions decided by chance before the experiment begins. Even if the neonate is born with an innate cross modal matching of the face they see and the motor actions they must do to achieve the expression, engaging with other faces is not indicative of intersubjectivity. The premise behind claiming intersubjectivity is that the face they are imitating is reacting to their expressions. There can be no intersubjectivity between an infant and a picture of an expressing face as the infant would not be regarding the face as reactive, let alone intersubjective.

In a comprehensive review of the literature that studies neonatal imitation, Anisfeld (1991, 1996) found that tongue protrusion (TP) was the only expression that consistently lead to matching behaviour at various points in infants between 2 days to 7 weeks. Out of 23 separate experiments with TP models only 12 demonstrated a significant increase in the rates of TP imitation, for all other expressions and gestures the results were inconclusive (Anisfeld, 1991, 1996).

Jones (1996) proposes an alternative explanation to the observed imitation of tongue protrusion claiming that many stimuli lead to TP. In a review of the presence of TP in related studies Jones discovered that TP is produced when an infant is aroused by various stimuli; including 'a touch on the palm, the sight of a looming and receding black pen, the sight of a box with a bright blue lining opening and closing and randomly blinking coloured lights' (Jones, 1996). This indicates that TP could be explained as blanket exploratory behaviour. In a study of gaze, infants look longer at TP in adults more than mouth opening (MO) so the prevalent evidence that infants consistently imitate TP is coincidental, the infant finds TP arousing and the way infants respond to stimulus just happens to also be by TP. They support this evidence with evidence that at the point that infants learn to control their arms and hands, learning to reach as exploratory behaviour at more than 20 weeks when the point at which imitation of TP in infants begins to decline. In short, Jones (1996) is arguing that the prevalence of imitation of

TP is just coincidental matching of interesting visual stimuli with infant's exploratory behaviour (Jones, 1996).

However, it would be difficult to disprove that the actions are not independent of one another as it could be that infants are imitating *and* exploring. Is TP is the only exploratory behaviour within their skill set?

Revisiting the evidence of neonate imitation, Jones (2006) contests that cross modal matching is an innate ability. Suggesting that mimicry is would be a complex and multidimensional capacity for a neonate:

1. 'there must be a motivation to imitate'
2. 'the infant must be able to identify the body part that produces the movement'
3. 'the infant must be able to identify the body on themselves'
4. 'the infant must be able to map the movement between the bodies- taking into account that their bodies are smaller so must produce a smaller version of the movement'
5. 'the infant must replicate the movement remembering the direction, shape and speed'
6. 'the infant would be required to control the muscles appropriately to complete the movement'

Addressing this Jones (2006) presents a study of infant-carer interactions where some of the stimulus expression is always accompanied with a vocal sound. This type of stimulus was produced in a circular fashion where carers imitate the infant's movements and reproduce them with a sound. Infants tend to emulate movements that are associated with sounds. If the sounds were taken away there was less of an imitative response, any imitation observed was usually previously well established between the infant and carer, so the carer may have imitated the infant in the first instance. In all studies the carer imitated the infant's actions more, also giving more smiles when the infant imitated the carer than when the infant produced a contingent response. The proprioceptive element of the imitation was achieved by the infant experimenting with different movements to match the outcome. Furthermore the infant's emulation didn't always match the same movement to achieve the outcome, for example in a reaching task the carer would reach and press a button in one fashion and the infant would explore different manners in which to press the button, not always using the same movement to achieve the

outcome.

This suggests that the outcome is more important than the mimicry behaviour itself. By redefining the matching purpose Jones (2006) suggests that infants learn to imitate by being imitated and this forms the basis of associative learning 'gaining knowledge of equivalencies between one's own body and movements and the bodies and movements of others' (Jones, 2006).

Similarly, Ray and Heyes (2011) argue against the so called 'poverty of stimulus' argument- that although imitation is more complex than a reflex it's performance is not based on learning, so therefore must be innate, a 'starting state nativism', a postulation that led to similar assumptions about innateness in Meltzoff and Brooks (2001). Alternatively, they claim that when infants engage in cross modal matching the imitations are straight in shape and do not stop but get better. As we have seen, TP is the only act that has recurrently been measured to be imitated by an infant dropped off after the infant could use their hands (Ray and Heyes, 2011).

Ray and Heyes (2011) suggest the 'wealth of stimulus' argument, suggesting that Associative Sequence Learning (ASL) accounts for the solution to the correspondence problem (the complexity of the act of mapping a visual perception cross modally to proprioceptive acts as outlined by Jones (1996)). ASL refers to the acquisition of equivalent experiences that are contiguous and contingent to the imitated act. In agreement with Jones (2006), Ray and Heyes (2011) claim that because infants are frequently imitated by their carers there is ample 'opportunity for infants to concurrently see and do, do and see an action', their imitations are constantly rewarded by carers. Not only does this contribute to equivalent associations but the authors also found evidence that infants work so they can see their limbs for visual feedback (Ray and Heyes, 2011).

However, the 'wealth of stimulus' argument does not address the question of what motivates the neonate or infant to engage in mimicry acts in the first place? The authors suggest that imitation is culturally inherited behaviour, that infants do it because they learn to do it as their carers constantly imitate them. This is a compelling argument given the recent evidence from Simpson et al (2014) that cross culturally there is a difference in levels of imitation between infant and mothers from different cultures. North American and European mothers tend to

imitate their infants, this leads to more imitation from the infant in contrast to other cultures. Infants and mothers from Japanese and rural African cultures tend to respond with contiguous acts without imitation, and whose young produce comparatively less imitation (Simpson et al., 2014). This idea still claims an imitative aspect to behaviour in a broad sense in that infant's behaviour matches their carer in whichever way, congruently imitative or non-imitatively but contiguous acts.

Condon and Sander (1974) conducted a study using the technique of meticulously nudging 16mm films of infant-carer interactions along a timeline to observe synchrony at a micro-kinesic level between a neonate's movement in response to adult speech when the infant was not looking at the adult. The adult's speech is segmented into time units based on phones, syllables and words. The neonate's body movements were segmented by units of sustained movement in a particular direction lasting usually from 0.04 to 0.16 seconds. This included all body parts, brows, eyes and mouth- these units bounded by changes in direction of that particular body part or termination of the moment altogether. The authors claim to have consistently observed temporal synchrony or entrainment between the infant's movement units and the adult's vocal segments. Suggesting that the infant performs expressions of participation within shared organisational forms rather than limited to sending isolated and discrete messages (Condon and Sander, 1974). There are criticisms of this particular methodology because the forms of 'shared organisation forms' can be indeterminately attributed in meaningful ways to any movement the experimenter deems appropriate.

Trevarthen (1998) used the same technique focussing on rhythmic qualities and coordination. Observing that these interactions used rhythmicity as a common timebase- a way to identify correspondences and comment on them within interactional exchanges. Trevarthen observed a precise coordination in (to 10ms) that he regarded as 'intrinsic rhythmicity' they use to base a 'primitive perception of agency, causality and intentionality'. Both infant and carer reciprocally engage in expressions that are contingent to the sequence within the interaction, entraining integrated expressive movement through temporal and morphological markers such as timing, form and intensity. This resonation within interactions between infant and carer termed as 'proto conversations' by Bateson (1971) where the contingent expression described above forms a mutual narrative of experience, each anticipating the other's commentaries moment to moment (Bateson, 1971). Here the infant can actively experience their carers experience of themselves

and this allows for imitative responses to occur at a moment in the stream of interaction as: affirmations, acceptances or commentaries, to accentuated displays of the other based on emotionally appropriate behaviour rather than straight imitation. Trevarthen claims that the infant motivation to actively engage with sympathetic other is purposeful and goes beyond attracting care for biological needs (a result of being born highly dependent) but are used to form emotional attachment to others in which to purposefully express motives and emotions with and for the sake of learning (and practicing) normative cultural / social expressions and conventions. Unlike Meltzoff etc, Trevarthen regards the mimicry behaviour between infant and carer more like an entrainment, using cross modal (multimodal) modified patterns and similar rhythms of movements that demonstrate moment to moment awareness of a shared narratives of feeling that is jointly attended to. As previously seen in Section 2.4, this is similar to Schegloff's emphasis on the strategic use of multi-modal resources as part of turn-taking sequences to defend intersubjectivity (Schegloff, 1992). Trevarthen suggests that new-borns possess an interpersonal theory of intersubjective intelligence that is more than imitation but is more dyadic or interactive in form, claiming infants are born with an innate conversational mind that goes beyond cross modal matching (Trevarthen, 1998).

Trevarthen and Aitken (2001) suggest that the dialogical exchange of affective states or referential narratives can be achieved through *bodily mimesis*. Often cross-modal, bodily mimesis is usually composed of modified patterns (for example, with similar rhythms of movement) of emotionally appropriate behaviour:

> These behaviours are not simply imitations of the forms of procedures or mannerisms- they are significant, emotionally charged interpersonal messages or displays, as, indeed all infant's imitations tend to be in some measures from the earliest days (Trevarthen and Aitken, 2001, p.14).

Trevarthen (2001) observed that there was a mutual regulation of interest's and feelings between couplets in 'intricate rhythmic patterns' in the expression of multimodal imitations of behaviour. So the imitative reciprocal responses are carefully timed in interaction sequences in order to 'act affirmations, acceptances, or commentaries' with respect to 'accentual displays of the other person'. Trevarthen concludes that the infant's behaviour has a coherent organisation that specifies the timing and form of body movements. There clearly is a two-way mirroring of the emotional values of expression and it is intersubjective at a fundamental level. Trevarthen states

that 'this organisation can react with appropriate dynamic changes to another person's dynamic expressions, matching their rhythms and accents' (Trevarthen and Aitken, 2001).

Daniel Sterns' *Interpersonal World of the Infant* (1985), argues that conceptualisation and reasoning are 'grounded in recurring patterns of our embodied interactions that can be characterised by cross-modal schemas such as containment, source-path-goal, compulsive force, attraction, scalarity and so on' (Stern, 1985, p. 93).

On intersubjectivity Stern remarks 'interactors' perception-behaviour links are coupled and interlaced with each other' and that this leads to a synchronous co-variation of movements from each recipient in response to each other. Stern proposes 'social understanding emerges from a dynamical process of interaction and coordination of two embodied subjects coupled together' (Stern, 1985, p. 94).

Stern comments on the infant's innate capacity in their ability to perform cross-modal matching that skilfully allow the transfer, amodally, the properties of duration, beat, rhythm. He continues to say: 'Interactions take on the form of organic rhythms and feelings of force, that is, the vitality affect of persons' (Stern, 1985).

So what of vitality affects and intersubjectivity? Stern thinks that for the constitution of intersubjectivity to occur, there must be a shared framework of meaning via a shared system of communication such as gesture, posture, or facial expression. Interaffectivity, the 'state mirroring and reciprocal cross-modal imitations of vitality affects', described by Stern in three stages of intersubjectivity.

- The detection of feeling state from overt behaviour (call)
- The response, or a cross-modal reciprocity contour (affect attunement) (response)
- The detection of this imitation or response as having to do with the original feeling experience (intersubjectivity confirmation)

Stern defines feeling qualities transmitted with distinctive activation contours which are captured by such kinetic terms as 'crescendo', 'decrescendo', 'fading', 'exploding', 'bursting', 'elongated', 'fleeting', 'pulsing', 'wavering' and so on. Always strongly rhythmic 'intermodal fluency' and synchronous 'affect attunement'. They give communicative meaning to the discrete action patterns shown representing categorical affects of anger, joy, fear and surprise etc. These be-

haviours are not simply imitations of the forms of procedures or mannerisms but they are 'significant, emotionally charged interpersonal messages or displays'. Attunement takes the experience of emotional resonance and recasts that experience into *another* form of expression.

Affect attunement is the performance of behaviours expressing the quality of feeling of a shared state, based on specific forms of intensity, time and shape of an expression, without imitating exact behaviour expression. It is the embedded and sequential attunement in this sequence that forms the quality of the relationship. In observations of infant/carer interactions, Stern finds evidence affect attunement occurs nonconsciously, and misattunements have the capacity to unsettle the child.

> Tracking and attuning with vitality affects permit one human to 'be with' another in the sense of sharing likely inner experiences of feeling-connectedness, of being in attunement with another. The analogous translation from perception of another person's behaviour to feelings involves the transmutation from the perception of timing, intensity, and shape via cross-modal fluency into felt vitality affects in ourselves (Stern, 1985, p.157).

Interaffectivity then, is the most immediate, closest to emergent experience form of sharing subjective experiences; a 'shareable inner universe' as a reciprocal interactional process across time where the focus of attention is known by both parties to be shared (Stern, 1985).

3.1.3 Summary

In summary, although neonate imitation is a widely accepted phenomenon there is a large amount of inconsistent data and where positive data was acquired a range of alternative explanations other than the idea that MNS governs the infant's behaviour are offered, the field of child development seemingly has moved on from the theory that neonate imitation is a simple automatic reproduction of perceived movements and most authors agree that there is more cognitive processing involved. A perception-behaviour link claimed to be based on a MNS would be present and observable from birth due to it's innateness. Even Meltzoff, the pioneer of neonatal imitation as a result of a simulative 'like me' cross modal correspondence, accepts that mimicry behaviour is a part of a higher goal based system.

Authors like Trevarthen and Stern propose a more complex, integrated and interactive occurring between infant and carer, forming imitative correspondences that could be related to the MNS. Even in this literature there is still a dynamic sense of 'moving' together to mutually understand each other's emotional or motive states, infants commenting on their perceptions multi-modally expressing inner qualities in actions that resonate with the way they experience it. Without a symbolic representation to communicate with, a neonate learns through mutual expression of equivalences between them and their carer. Trevarthen claims that this type of joint attention leads to communication influencing the infant's perception.

Beebe et al. (2003) compare Meltzoff, Trevarthen and Stern's accounts to matching behaviour or self/other correspondences in infant carer interactions. The authors note that each agree that the infant's capacity for cross-modal matching is key to detecting correspondences, and that infants have an innate, complex, pre-symbolic representational intelligence that is motivated, purposeful and acted out in social interaction. However, the definition of these matching behaviours or correspondences is different between the three authors. Meltzoff defines a one way static imitation model that is defined by form whereas Trevarthen and Stern use a more dynamic definition that considers timing, form and intensity that is more of a reciprocal interactional process. Returning to the MNS system, the mapping of visually perceived behaviour of another onto one's own motor plan or image, the authors discussed in this section disagree on how this system would work within these differing theories of correspondences (Beebe et al., 2003).

By using an adapted form of the Other's expressions, a more strategic display of understanding is given. Simply mimicking an expression cannot produce intersubjectivity because enacting a straight replication of an expression only acknowledges the expression itself without showing an understanding of the meaning behind the expression. By elaborating on an expression, producing a variation as a form of acceptance or commentary such as Trevarthen reports above, displays an understanding of the meaning behind the expression. In these simple affect expression couplets a quality of the feeling expressed is 'paraphrased' to demonstrate understanding.

3.2 Empirical studies on intersubjectivity in adult interactions

3.2.1 Behavioural mimicry

The claim that there is an automatic and nonconscious tendency to mimic one another in social interactions has been widely studied. Behavioural mimicry has been observed between interlocutors in self-adaptors or mannerisms such as foot-shaking and face-touching (Chartrand and Bargh, 1999; Yabar et al., 2006; Stel and Vonk, 2010; Lakin and Chartrand, 2003; Lakin et al., 2008), posture (Scheflen, 1964, 1972; LaFrance, 1979; Trout and Rosenfeld, 1980; Maurer and Tindall, 1983) more recently (Feese et al., 2011; Tia et al., 2011; Hagad et al., 2011), facial expression (Wallbott, 1991; Hatfield et al., 1994; Hess and Blairy, 2001), gesture (Kimbara, 2006; Parrill and Kimbara, 2006; Kimbara, 2008; Holler and Wilkin, 2011b) and yawning (Provine, 2005; Helt et al., 2010) among other motor movements (Chartrand and Lakin, 2013).

3.2.2 The perception-behaviour link

Previously reviewed in Section 2.3.1 are numerous studies of facial mimicry that make a case that the imitation of facial expression is the basis for emotional contagion and emotion recognition in others. Along with higher cognitive resources that allow for perspective taking (see Section 2.3.2), the experience of empathy is based on these mechanisms. This section details a comparable mechanism based on bodily forms of mimicry as opposed to facial mimicry. Although the main consequence within interaction of the mechanism is a prosocial effect, it is also claimed to be a basis for empathy.

Chartrand and Bargh (1999) define the *perception-behaviour link*: a nonconscious connection between the act of perceiving and the act of behaving, such that 'people automatically behave as they perceive' (Chartrand and Bargh, 1999; Lakin et al., 2006; Chartrand and van Baaren, 2009; Chartrand and Lakin, 2013). Drawing from Carpenter (1874) and James (1890) principle of *Ideomotor-action* that ties thinking with action; merely thinking about behaviour increases the tendency to engage in that behaviour. Jeannerod (1999a,b) go further to demonstrate that mentally simulating (imagining) an action activates the same premotor cortex neurons as performing the same action. They suggest a two-step process:

- automatic perceptual categorisation and interpretation of social behaviour (environment

to perception)

- perceptual activation continuing on to activate corresponding behaviour representation (perception to behaviour)

The sequence from environment to behaviour is said to occur nonconsciously and automatically. For this reason it would be expected that mimicked behaviour would follow any variability in the behaviour of an interactional partner. In an experiment they asked participants to describe photographs to two separate confederates, known to them as other participants, in two sessions. The confederates alternated behaviours in each session, one confederate would shake their foot and the other rub their face repeatedly during the interaction. Their results showed that these behaviours are indeed mimicked by the participant accordingly and demonstrate evidence of what they termed the *chameleon effect*. This effect occurred in minimal conditions: among strangers with no goal to affiliate with each other.

A second experiment investigated the prosocial effect of behavioural mimicry. Participants again described photographs to a confederate known to them as another participant. However, in this experiment there was only one confederate who alternated between mimicking the participants behaviour or remaining neutral. The participant was asked to rate the smoothness of the interaction and the likability of their interactional partner (the confederate). Participants reported that the interactions were smoother and their interactional partner more likeable in the mimicry condition.

Here the chameleon effect is described as a 'cause of interpersonal rapport and empathy'. The authors hypothesise that the perception-behaviour link that they claim as the mechanism behind behavioural mimicry is related to perspective-taking, the cognitive component of empathy-but not the affective component, as the passive and automatic mechanism does not rely on an emotional state. A third experiment also saw participants describe photographs to a confederate known to them as another participant. In this experiment the confederate both shook their foot and rubbed their face throughout the interaction. The participant then completed a questionnaire that rated their individual degree of perspective-taking and level of empathic concern. Their results show that participants that scored a higher degree of perspective taking displayed more mimicry than those who scored lower, however the level of empathic concern had no effect on mimicry. From these findings it was suggested that the imitation of postures

and gestures are a continual source of information throughout a social interaction, communicating messages indicating understanding and attention, leading to smoother interactions when mimicry was present (Chartrand and Bargh, 1999).

Chartrand and van Baaren (2009) propose two cognitive theories to explain the perception-behaviour link: Firstly Prinz (1990) *common coding principle* states that perceived events and planned actions share a common representational cognitive domain. For example, when perceiving a particular movement, such as observing someone tapping a finger, can induce or interfere when concurrently performing an action that shares the same features. Here, overlapping sensory and motor codes (where perception shares features with planned actions) results in a faster response time if the perceived action is similar to the planned action, say tapping an index finger whilst observing a tapping index finger. However when observing a finger on the opposite hand tapping will interfere with concurrent action as the same codes for the incompatible perceived events are activated at the same time as planned actions. Chartrand and van Baaren (2009) propose that 'the representations of action automatically lead to actual behaviour after a certain threshold', they however do not elaborate on how this 'threshold' operates within this mechanism. Brass et al. (2001) show further evidence that similarity in features of perception induce or interfere with planned actions by demonstrating a response time paradigm demonstrating that observing a certain finger movement facilitates the same action interferes with performing a different finger movement. Hesslow (2002) reflect this by making the claim that 'thinking is essentially a covert or 'weak' form of behaviour', simulation of action is suppressed behaviour and again this is because as the sensory domain that governs perception is linked to the motor structures that govern action. A second cognitive theory claims that perception and action share overlapping semantic 'schemas' (Barresi and Moore, 1996). These schemas impose a different structure on first person (imagined) actions and third person (perceived) actions but as a result of these overlapping representations, the two types of schemas can be active at the same time. Knoblich and Flach (2001) suggests that this link between perception and action is for the purpose of predicting the consequences of observed actions. The consequences of this overlap manifest in 'perception leading to action, and action also leads to perception- that is to interpret behaviour in a certain way'. The latter interpretation is why the consequences and performance of mimicry are bidirectional.

3.2.3 Mimesis and cognitive neuroscience

3.2.3.1 The mirror neuron system

Further developments in the field of cognitive neuroscience propose evidence for the neural correlate of these simulative theories in the brain. Much of this research is based on the discovery of *mirror neurons* in the brains of macaque monkeys in the early 1990s. By placing electrodes directly into the neural region claimed to code goal-related motor acts such as hand and mouth grasping, individual neurons were found to activate when the monkey executed an action *and* on the observation of the same action executed by another monkey. This discovery demonstrated a functional neural link between perception and action in the monkey's brain (Di Pellegrino et al., 1992; Rizzolatti et al., 1996; Gallese et al., 1996).

These original studies on macaque monkeys lead to numerous further experimental studies to test whether a similar mirror mechanism is present in humans. These studies on humans use non-invasive techniques such as the aforementioned EEG scans[1] or more frequently fMRI scans[2] operate at a vastly larger scale- measuring neural activity in small brain regions as opposed to measuring single-neuron activity as the invasive electrophysiological experiments on macaque monkeys do. For this reason these studies are not able to directly detect individual mirror neurons but instead detect neural regions that are active when the person executes an action and also active when the person perceives another individual executing the same action (Keysers and Gazzola, 2010). From these findings it is proposed that these neural regions contain mirror neurons and constitute to the human equivalent of a mirror neuron system (Iacoboni et al., 1999; Gallese et al., 2001; Decety and Chaminade, 2003a,b; Pobric and De c. Hamilton, 2006; Saygin, 2007; Gazzola and Keysers, 2009; Tranel et al., 2010).

There have been a number of arguments that question whether the mirror neuron paradigm can be extended from studies on macaque monkeys to studies on humans because of the lack of systematically comparable measures but also the expectation that the mirror neuron paradigm in macaque monkeys can be translatable to human brains is argued to be an unreliable starting point to base such a body of research on (Pascolo and Budai, 2013). Furthermore, brain

[1] See section 3.1.1

[2] functional magnetic resonance imaging (fMRI) creates three-dimensional models of the brain using a strong magnetic field and radio waves to show blood flow in the brain to detect areas of activity

imaging techniques used on humans cannot directly measure mirror neuron activity as the fMRI responses generated by neighbouring visual, motor and visuo-motor neurons do not fall into the same class of neuron as mirror neurons (Dinstein et al., 2008).

One study used invasive electrophysiological measures to positively detect mirror neurons in the brains of 21 patients under treatment for intractable epilepsy. Mukamel et al. (2010) found a small number of neurons were found to fire on perception and execution of action (such as hand grasping actions, and facial emotional expressions). As the electrodes used to measure neuronal activity are placed for clinical purposes, the neurons detected were not located in the premotor cortex as in original studies on macaque monkeys. Instead, they were sited in supplementary motor regions - suggesting the presence of mirror neurons in multiple regions in the brain. fMRI responses suggest compatible activity of a mirror mechanism in these regions (Gazzola and Keysers, 2009).

In addition a portion of neurons that exhibited 'anti-mirror' properties were observed. These neurons demonstrated increased activity during execution of an action but inhibition (activity below baseline) during action perception. It is suggested if actions executed by others are represented in the brain's motor system- neurons that respond with an opposite pattern of excitation could be the basis of a control mechanism that sees that perceived actions are not always imitated. It is speculated that this could explain compulsory imitative behaviour in neurological patients with extensive brain lesions (Mukamel et al., 2010). This is also relevant to the observation that macaque monkeys do not engage in imitative behaviour (Wilson and Knoblich, 2005). Keysers and Gazzola (2010) suggest 'anti-mirror' neurons could also be how our own actions are disambiguated from the actions of others:

> In combination with mirror neurons, they could help the brain perform an inner simulation of other people's actions while at the same time selectively blocking overt motor output and disambiguate who performed the action (Keysers and Gazzola, 2010, p. 354).

Pascolo and Budai (2013) point out that much of the fMRI studies on mirror mechanism and Mukamel et als. (2010) above study do not fully take in account the temporal correlation between perception of actions and corresponding neuronal activity. As fMRI imaging requires time to obtain useful data there are potential discrepancies as these studies may not take into

account temporal delays in activation detection. Furthermore, many of the fMRI and PET studies only involve the passive observation of an action by participants, comparing the area neural activity is detected in with areas that are inferred to be involved in it's execution. Turella et al. (2009) highlights that often the areas involved in observation and execution of an action did not match. Connolly (2012) also disputes the significance of neuro-imaging studies that rely on inferences on brain regions: 'Those observations tell us what zones of the body/brain system are excited, but they do not specify the contents'.

Originally the *direct matching mechanism hypothesis* was proposed, stating that mirror neurons are a type of visuo-motor neurons that serve to match visual presentations of action onto a corresponding motor representations of action in the brain, claiming existing motor knowledge of an action is drawn upon to understand and recognise perceived actions. In this sense, the mirror neuron system is interpreted as a direct mechanism of action understanding (Gallese et al., 2001).

However the direct matching mechanism hypothesis has now been heavily disputed, It has already been observed that the mirror neuron system does not respond to perceived actions by their visual description alone but on the anticipation of end goal (Gallese, 2008). As the mirror neuron system is activated *only* when the perceived action is goal-orientated (these include pantomimes of goal-orientated action, object-orientated actions and gestures with communicative goal), this would require the goal/intention of the action to have already have been recognised and understood prior to any activation of the mirror neuron system. An action would have to be understood first for the mirror neuron system to be activated at all (Kosonogov, 2012). Although this does not invalidate the claim that the mirror neuron system can be drawn upon to predict *subsequent* actions (Fogassi et al., 2005). Actions are heavily contextualised such that there can be multiple different goals for the same action depending on situation. As the mirror neuron system responds more strongly to perceived actions embedded in particular context than to the same action but without context, this suggests that the mirror neuron system is more likely related to the intention behind perceived actions (Iacoboni et al., 2005). In addition, Chaminade et al. (2002) conducted an experiment testing whether imitating a gesture activates the neural processing of the intention (or goal) underlying the observed action. The study used Positron emission tomograph (PET) scans[3] of participants observing hand gestures

[3]A functional imaging technique used to detect metabolic processes in the body

that move Lego bricks from one place to another. The participant either imitated the exact movement to move the Lego brick from one place to another or freely completed the action in whichever way suited. They found that despite whether the action was imitated of performed freely neural activity was just as prominent.

Furthermore, studies have shown that the mirror neuron system is activated when a goal-orientated action is just heard (but not seen) both in monkeys (Kohler et al., 2002; Keysers et al., 2003; Pineda, 2005) and in humans (Gazzola et al., 2006; Etzel et al., 2008), as well as when read from a written description of an action (Aziz-Zadeh et al., 2006). In addition, it has been shown that the mirror neuron system still activates even if part of the action is hidden but the context is already known, such as reaching for objects out of present view but previously seen to be there (Umilta et al., 2001; Fogassi et al., 2005). Another fMRI study showed that by even passively listening to nonverbal emotive expressions, such as laughing or crying, cross-modally activates the neural areas involved in the control of the associated facial motor movement (Warren et al., 2006). This suggests that actions are understood by higher cognitive processes first without the need to simulate action by mapping a visual presentation of an action to a motor representation using the mirror neuron system. For these reasons it is deemed implausible that the function of the mirror neuron system is to understand and recognise perceived actions.

Is there a cognitive function to the mirror neuron system in retrieving the motor representation on perception of a goal-orientated actions? Kosonogov (2012) points out that the mirror neuron system is not activated for actions not worth repeating; suggesting that the function of the mirror neuron system could be to implicitly reproduce actions so a later imitation of the action can be facilitated *if* needed:

> If the action is important for the observer and can be useful in his own motor repertoire, his/her mirror neuron system implicitly reproduces the action, retrieving the kinematics and sensory consequences the observer experienced in the past while executing the same action (Kosonogov, 2012, p. 499).

Wilson and Knoblich (2005) highlight two purposes behind imitation facilitation that have emerged: for observational learning and to facilitate social behaviour. They make the claim that the mirror neuron system allows for a motor activation to 'contribute' to the interpretation

of another's behaviour by acting as an emulator for perceived actions, the intention behind an action is inferred by 'reinstating sensory consequences of an action' and cognitively building on this emulation to understand and predict other's behaviour.

3.2.3.2 Associative learning and mirror responses

Heyes (2001) attempts to explain *how* the mirror neuron system can match perceived and executed actions with a theory of Associative Sequence Learning (ASL). This claims that because the perception and execution of an action are paired sensorimotor experiences there forms a 'bidirectional associative link' between them. For example, the visual stimulus of perceiving one's hand reach to grab an object is associated with the motor program that allow one to execute the action. It is this associative link that leads to the phenomena of *automatic imitation*, the reason that 'humans are prone, in an unwilled and unreasoned way, to copy the actions of others' (Heyes, 2010a, p. 1).

Often explained in terms of *Hebbian learning*: that brain organisation is such that simultaneous (or temporally consecutive) firing of neurons builds up 'synaptic strength' between neurons- after time activation of one neuron enhances the efficacy of the other, facilitating it's firing Hebb (1949). Keysers and Perrett (2004) relate this to the mirror neuron system by suggesting that synapses connecting sensory and motor representations of an action are strengthened as a result of the 'temporal precedence' between sensory stimuli and the motor programs that govern executing actions. This is compatible with findings that mirror neuron system is also activated on an auditory stimulus of an action. The neuronal correlates of perceiving and executing an action are increasingly inter-associated. Mirror neurons, in this sense, can be more accurately described as multimodal association neurons (Keysers, 2009; Hickok, 2009).

As previously mentioned, ASL accounts for the correspondence problem- for the cross-modal problem stating that in some situations we cannot perceive the facial or bodily expression of our own actions so it is impossible to employ self-observation in order to form associative links between sensory and motor representations. Heyes (2001) proposes other sources that sensorimotor links can be formed- mirror reflections, imitation by carers in infancy (see Section 3.1.2) and perceiving socially concurrent expressions or movements from others in response to a shared stimulus. As macaque monkeys have been found not to engage in imitative behaviour (Roth

and Dicke, 2005; Lyons et al., 2006), or watch themselves in mirrors in their natural habitat, ASL theory can only be applicable for actions that can be directly self-observed by the monkey[4]. This could explain how the discovery of mirror neurons in monkeys are limited to hand grasping behaviours (Di Pellegrino et al., 1992; Rizzolatti et al., 1996).

Theoretically, the ASL model is consistent with the absence of reliable evidence of innate behavioural mimicry (see Section 3.1), at this neonatal stage infants are yet to have opportunity long-term associations between sensorimotor stimulus, suggesting that these are established over the infant's development. Hereby, if imitative responses are established through an associative learning process, these should be reversible through training. Haggard et al. (2005) find evidence that the tendency to make imitative responses can be inhibited with training. In their experiments subjects are either shown an image of an open hand or a hand closed into a fist, subjects were instructed to either imitate the hand configuration or perform the opposite hand configuration. Their study shows that performance response times for perceived images that are compatible with the instructed action are typically quicker, but do improve with training. They suggest that inhibitory links or new sensorimotor associations can be formed when the perceived action is incompatible with the action they are instructed to perform, somewhat overriding the original imitative effect. These findings are supported in a similar study on compatible and incompatible finger abductions, demonstrating a similar reversal of the mirror effect with training (Catmur et al., 2007). Note, these studies are not experiments investigating behavioural mimicry as the participants are instructed on what behaviour to perform and the measures are revealed through response times.

Heyes (2010b) proposes that long-term associations that link the topographical features of sensory representations with corresponding motor representations of action should see 'experts' in these actions to have established associations- resulting in stronger mirror responses. This is supported by a study involving experienced ballet dancers- indicating that activity of the mirror neuron system is modulated by experience, showing stronger responses for known movements (Calvo-Merino et al., 2005).

What Heyes (2010a) terms *automatic imitation* is different to the types of behavioural mimicry as discussed by Chartrand and Bargh (1999) and subsequent work on the chameleon effect,

[4]With the exception of mouth actions. However see Paukner et al. (2009), although this only provides a study on Capuchin monkeys it could be indicative of how experimenters imitate monkeys generally in this circumstance

as it refers to the tendency to imitate in compatible/incompatible stimulus studies measured dependent on time responses. The ASL model is nonetheless regarded as consistent with the chameleon effect and behavioural mimicry deemed a byproduct of the sensorimotor links established through associative learning. However, in contrast to conventional theories that describe the mirror neuron system, associative learning can account for the tendency toward *complementary* action in social situations. Newman-Norland et al. (2007) found stronger activity in the mirror neuron system when instructed to perform complementary action- grabbing a beer mug by the handle when it is passed by the mug, than when instructed to perform imitative action- to grab the mug itself. This also avoids the problem that Pascolo and Budai (2013) set out, that automatic imitation wouldn't help a boxer when anticipating a response to incoming attacks- which quickly require complementary, defensive, action as the ASL model depends on the motor responses associated with the perceptual stimuli. Tiedens and Fragale (2003) showed that when confederates adopted dominant or submissive postural stances, participants responded with the complementary postural stances and not imitative stances.

This applies to the complementary nature of social interaction 'a critical component of any social interaction is a complementary, contingent oscillation between social roles' experiment EEG scanning alternating turns and imitating a models gestures partner on screen told to imitate alternative (Nadel and Dumas, 2014)

Furthermore, the associative learning model has been applied more generally, including situations where a stimulus can elicit a dual response of learnt associations. Such as the automatic imitation described here and spatial compatibility responses (responses that depend on the position the movement relative to an external frame of reference) (Heyes et al., 2013), or automatic imitation and spatial perspective taking (cognitively rotating one's own frame of reference rather than rotating objects to emulate another spatial perspective) (Kessler and Thomson, 2010). Both studies report that spatial responses are stronger than imitative ones, these are still deemed as responses borne out of long term associative learning. It is not known if there are inhibitory links between dual-route architectures such as these. It has also been noted that imitative responses are facilitated by an on-looker - called the 'audience effect' (Wang and Hamilton, 2014).

A number of more complex variations of the ASL model suggests that features that make up the perceptual representation, motor program and contextual information relating to an action

is neurally bound in a network of associated neurons, called 'event files' or 'feature maps'. The bidirectional link between the networks of neurons causes the activation of one feature to retrieve one or more features within the network. Described as bound distributed representations that are the result of entrenched patterns of activation of sensory modalities and the motor system (Barsalou et al., 2003; Hommel, 2004).

3.2.3.3 Empathy, embodied simulation and the mirror system

Gallese and Goldman (1998) suggests that the mirror neuron system is the underpinning mechanism of 'mind-reading' as it allows us to simulatively model other's actions goals to contribute to an understanding of other's mental states. The 'motor resonances' that other people's actions trigger in us contributes to the ascription of intentions behind observed actions that are assumed to underpin their behaviour. Favouring a simulative stance on a theory of mind, Gallese (2001) proposes a 'shared manifold of intersubjectivity', as the capacity for understanding others as intentional agents is established in the relational nature of action. Sinigaglia and Carrado (2011) describes a 'reciprocity of being', that we are implicitly understand others as similar because we share a common embodiment. This extends to the neural substrate, when actions are executed or affective experiences are had, a reciprocal activation of shared representations are experienced by others.

Gallese (2008) suggests that our capacity to empathise with others is mediated by embodied simulation mechanisms, that is, 'by the activation of the same neural circuits underpinning our own emotional and sensory experiences'. In agreement with Hatfield et al. (1994) ideas on emotional contagion (see section 2.3.1), Gallese (2008) also claims that the mirroring mechanism seems to be involved with our capacity to converge emotionally. 'When perceiving others expressing a given basic emotion by means of their facial mimicry, the observer's facial muscles activate in a congruent manner, with an intensity proportional to their empathic nature.' Gallese and Lakoff (2005) notes that as the mirror matching mechanism represents content independently of a self/other distinction- Gallese (2008) describe a 'we-centric' interpersonal link is formed on the basis of this imitation mechanism, and influences the capacity to share meanings, intentions and emotions with others and therefore grounding our experiences. Gallese (2008) claims when imitating the facial expression of basic emotions, the same restricted group of brain structures are activated (Gallese, 2008). However, Jacob and Jeannerod (2005) point

out the problem with the MNS is that it is for motor simulation but not social simulation, or rather, the same action can be used for different goals. They emphasise that this is the reason social intention cannot be determined by simulating observed movements, but that it requires context for an intersubjective understanding.

Decety and Sommerville (2003); Grèzes and Julie (2006) propose the shared representations model. Claiming that we may understand the actions and emotions of others in terms of our own motor system. For example, while observing a smile, the observer would mimic same facial muscles at a sub-threshold level and this would activate the same neural network in the observer and cause them to feel happiness. This simulation is a conscious reactivation of previously executed actions stored in memory- a direct link between perception and action prompts observer to resonate via recalled mental states with the emotional state of another. This is neurally different from one's own experience- subjective experiences activate the anterior cortex whereas inwardly simulating another's experiences activates the posterior cortex, although it is a common circuit, this difference is claimed to be qualitatively different- allowing for a distinction between empathic responses from our own sensations.

Sinigaglia and Carrado (2011); Gallese (2013) later propose a mirror mechanism driven by 'embodied simulation'- As representations of action and affect are bodily in format one can reuse their own mental states or processes to functionally attribute mental states to others. Chaminade et al. (2010) describe embodied simulation mechanism as *resonance*- 'the mechanism by which the neural substrates involved in the internal representation of actions, as well as emotions and sensations, are also recruited when perceiving another individual experiencing the same action, emotion or sensation'.

Singer (2006); Singer and Lamm (2009) describe the affective side as a bottom-up process activated by a perception-behaviour link, the cognitive side as a top-down process which includes a contextual appraisal and cognitive perspective taking. These are separate neural circuits with different developmental pathways. The cognitive process modulates the representations produced by the automatic affective process, but also serves in the absence of bottom-up stimulation, from imaginative perspective-taking.

3.2.3.4 Empathy for pain and neural correlates

A general issue with analysing the neural activity of affective states is their lack of contextual specificity. Bastiaansen et al. (2009) reviews literature on emotion / affective dimension and the MNS, pointing out that emotions are too non-specific- affective states without context are hard to simulate in experiments and there is no concrete validation that the exact regions or structures can be found. A neural response to pain and neural activity generated from the observation of another in pain, on the other hand, is similarly caused in participants and more acutely based in the body.

Several neuroimaging studies have shown documentation of a common neural substrate involved in representing other's and our own pain, indicating that pain is an affective dimension that has significant evidence of a neural mirror mechanism (Lamm et al., 2011).

Jackson et al. (2006a) state that the representation framework of subjective experience of pain includes somatic-sensory features (Nociception from nerve receptors- the sensation of pain) and the affective features of pain, this framework is termed the *pain matrix* (Singer et al., 2004). Jackson et al. (2005) showed participants photographs of painful events, an fMRI scan detected activation in the anterior cingulate cortex (ACC) and the anterior insula, neural areas known for their role in affective pain processing, but did not detect activation in somatic sensory cortices related to specific body parts. It is suggested that the perception of pain in others to an unspecific state of arousal (likened to distress or anxiety). Jackson et al. (2006b) ran further fMRI studies asking participants to imagine themselves in pain and imagine another in pain. The resulting scans revealed that the imagined self-perspective of pain activated the somatosensory cortex despite there being no physical sensory stimulus of pain. It is suggested that the non-overlapping part of the pain matrix has a role in our sense of agency and self-identification.

Facial expressions of pain are said to be an important cue in the social communication of pain. Botvinick et al. (2005) used fMRI on participants exposed to a video sequence of facial expressions of pain - the same participants were subjected to painful thermal stimulation. Imaging found common neural substrates that activated on perceiving a facial expression of pain, and being subjected to pain. This was echoed in Morrison et al. (2004), a study that found the same results when both being subjected to a painful pinprick to fingertip, and observing another

person's hand being subjected to the same pain stimulation. In a similar study, Saarela and Hlushchuk (2007) found not only are the affective dimensions of pain activated when watching facial expression of pain, participants could determine the intensity of the pain, and this was correlated with the strength of the neural activation. However, Heath (1989) investigated how pain was expressed in doctor/patient consultations, findings suggesting pain behaviour or expression, such as wincing or crying out, were designed to be informative to pain assessment. Participants paid attention to the sensation only in order to express pain behaviour when relevant to the diagnosis, or elicited through eye contact with the doctor. Heath concluded the expression of suffering had a communicative function, suggesting there may not always be correspondence between the sensation of pain and expression of it. Singer et al. (2004) used fMRI to assess brain activity in participants when being subjected to pain, and observing a loved one in pain. When comparing the experience of being subjected to pain with observing a loved one in pain, neurologically the entire pain matrix was not activated - only cognitive representations within the affective dimension of pain were activated.

Osaka et al. (2004) finds activation in the same areas in fMRI scans when exposing participants to simply onomatopoeic words expressing affective pain, suggesting the stimulus activating the affective dimension can be abstract and broad.

Empathy for pain is underpinned by neural structures that are also drawn upon in a personal and direct experience of pain. Danziger et al. (2006) tested whether empathy for pain is still felt in the same way by people who are insensitive to pain, they found although there is an underestimation of severity there is still an affective response. Similarly to Singer (2006), it is suggested the neural correlates discussed here form bottom-up processes driven by stimuli and these are mixed with top-down processes (beliefs, knowledge and experience) to fully empathise with the pain in others. This study shows people who are insensitive to pain can make inferences from top-down processes.

3.2.4 The mimicry prosocial effect

Behavioural mimicry is often regarded as a nonconscious communicative strategy in a repertoire of prosocial behaviours. It is frequently related to empathy, rapport and liking. Early observations of posture in dyads (therapist/patient interactions) and groups (classroom settings)

predicted postural convergence is a consequence of shared viewpoints and rapport (Scheflen, 1964, 1972; LaFrance, 1982). Maurer and Tindall (1983) later reflected this prediction finding postural congruence positively correlated with rapport and empathy between adolescents and a career counsellor. Postural congruence is also viewed as a marker of solidarity with a group, signifying affiliation (Beattie, 1988) (See Section 3.2.6 for a review of literature on posture). As previously mentioned, behavioural mimicry has been found to positively influence liking serving as 'social glue' (Chartrand and Bargh, 1999; Lakin et al., 2003, 2006). Interpersonal synchrony has been shown to increase rapport and affiliation (Bernieri and Rosenthal, 1991; Hove and Risen, 2009).

> Imitated postural configurations are a source of information about an ongoing social interaction, as they communicated messages about liking and understanding. Adopting similar postures increases sense of involvedness, togetherness, being 'in step' and compatibility. (Chartrand and Bargh, 1999, p. 896).

Numerous prosocial effects of being mimicked within an interaction have also been reported, such as feelings of trust (Guéguen et al., 2013; Verberne et al., 2013), a perception of a just world (Stel et al., 2012), decrease in racial bias (Inzlicht et al., 2012), decrease in victim blaming (Stel et al., 2012), invoking the willingness to help others or engage in charitable behaviour (Van Baaren et al., 2004a; Stel et al., 2008), better attitude towards products (Tanner et al., 2008), better negotiating (Maddux et al., 2008) and the persuasiveness of a message (Van Swol, 2003; Bailenson and Yee, 2005; Bailenson et al., 2008). On a personal level, being mimicked has been show to enhance mood (van Baaren et al., 2006), self-esteem (Chartrand and van Baaren, 2009), learning (Zhou, 2012), recall of surroundings (Van Baaren et al., 2004b) and self-control (Finkel et al., 2006; Dalton et al., 2010).

van Baaren et al. (2009) suggested being mimicked can evoke a general sense of similarity. This is shown in a study where a participant is mimicked (or not) by a confederate during an interaction discussing recent advertisements, participants then rated similarity between two images. Mimicked participants rated the images as more similar than the participants who were not mimicked by the confederate. This experiment demonstrated when people are imitated they perceive more similarity between objects and feel more similar to the other. This process was a circular one, in the more mimicry between interlocutors, the more positive the consequences of the imitation are. The phenomenon is termed the 'mimicry-prosocial effect'- mimicry strength-

ens a sense of attunement and connectedness between interactors'. It is suggested the reason for this prosocial effect of mimicry lies in the assimilation of orientations (perception of objects), cognitive styles and mental states it engenders.

It should be noted that although these studies highlight many prosocial effects related to behavioural mimicry, these in themselves are not indicators that people are experiencing empathy for whomever they are interacting with. Feeling a sense of rapport or similarity to someone is not defined the same as an experience of empathy- understanding the experience of another by feeling their experience.

As mimicry and laughter exhibit functionally overlapping social signals of affiliation and prosocial effects, a study that analysed mimicry during laughing in interaction asked- are gestures mimicked more frequently during laughter? Looking specifically for mimicry of hand movements within 3 seconds of an original hand gesture; the results showed more gestures performed during laughing but there was no increase in mimicry of hand gestures during laughing than outside of laughing episodes, despite more increase in gesture at these points. They did find, however, that mimicry increased in the second half of the interaction (Griffin et al., 2015). Van Der Zee et al. (2015) ran a study detecting deception using full body motion capture data. Participants were required to tell the truth or lie in an interview. Full body motion capture detected that participants did not freeze their body movements while lying, overall more body movement was exhibited. This was indicative of 75% of lying instances.

As the perception-behaviour link is not influenced by a conscious choice or guidance, behavioural mimicry should occur even amongst strangers and in the absence of a reason to do so (Chartrand and Lakin, 2013). However, there are many social moderators with the documented ability to mediate the perception-behaviour links effects, these are detailed below. These indicate mimicry is drawn upon selectively depending on social, cultural and emotional context.

One begs the question - if the perception-behaviour link is a unitary process to facilitate mirror responses, how is this flexibility explained? Wang and Hamilton (2012) suggests a top-down model (STORM - Social top-down response modulation) of the process leads to automatic imitation. The model takes into account the selective, and varying degrees of mimicry appropriate in different situations. The signals produced from activation of the mirror neuron system are modulated by social information influence the mimicry response. Modulation comes from a

mentalising system evaluating recipient of the mimicry response, priming, contextual information and so on. However, it remains to be seen if there are any neural mechanisms forming the basis of this model.

3.2.4.1 Facilitators/Inhibitors

Goal to affiliate

Lakin et al. (2003, 2008) argue the link between mimicry and liking would lead to increased mimicry from subjects pursuing an affiliation goal.

An initial study involved either directly instructing participants to pursue an affiliation goal with a confederate or subliminally priming subjects with an affiliation goal by exposing them to prosocial words such as 'friend', 'affiliate' and so on.. All participants watched a videotape of a confederate complete mundane tasks, participants were told to remember the confederates behaviour. Both the explicit and subliminal priming of an affiliation goal engendered more mimicry of the confederate in the video than when there was no affiliation goal.

In a second study participants were again subliminally primed with an affiliation goal. The researchers deemed subliminal priming to be sufficient to incite an affiliation goal as there was no difference between the explicit or subliminal priming conditions in the above experiment, a separate study also shows subliminal priming with social and anti-social words influences automatic imitation (Leighton et al., 2010). The participants then interacted with chat-software where their interaction was designed to fail or succeed at the affiliation goal. Participants then interacted face-to-face with a confederate. They found participants who failed their affiliation goal at the prior online interaction produced more mimicry in the second face-to-face interaction. These results demonstrate pressure to form an affiliation in an interaction engenders more mimicry. Furthermore, these interactions were reported as smoother, with more rapport by both the participant and the confederate (Lakin et al., 2003).

An additional study tested how social exclusion from a group influences mimicry. The experimenters instructed participants to play an online ball game with three other players, the three other players either excluded or included the subject. Participants were then asked to interact with another participant (a confederate) describing photographs to one another. The confederate shook their foot during the interaction. The results were positive to their hypothe-

sis, excluded subjects nonconsciously mimic the confederates foot shaking more than included subjects. In addition, the confederate rated the interactions with the excluded subjects to be smoother. It is suggested nonconscious mimicry evolved to elicit more successful interactions in aid of social context such as exclusion or ostracism (Lakin et al., 2008). It has even been shown simply feeling different from others engenders more mimicry (Chartrand and van Baaren, 2009).

Conversely, further studies show people exhibit less mimicry when attempting to disaffiliate- such as when interacting with stigmatised individuals, for example people with facial scars or people who are obese (Johnston, 2002) and outgroup members (Yabar et al., 2006; Bourgeois and Hess, 2008).

Existing rapport

The link between rapport and mimicry is further supported by studies reporting participants mimic friends more than strangers, liked more than unliked confederates (Likowski et al., 2008; McIntosh, 2006), people with shared opinions than diverging (Chartrand and Lakin, 2013) and even confederates pre-determined as honest more than dishonest (Stel et al., 2010b). Yet, although mimicry has been shown to contribute to smoother interactions and better understanding between interlocutors, Stel et al. (2010a) reports mimicry doesn't improve liking of already disliked people.

Self construal

van Baaren et al. (2003); Ashton-James et al. (2007) report mimicry is influenced by one's sense of self-construal, people of a culture with an interdependent self-construal exhibit more mimicry than cultures with a more independent self-construal. Participants who are primed with words relating to an interdependent self-construal, such as 'together' or 'team', also exhibited more mimicry than participants of an independent self-construal. Suggesting a 'complex relationship between culture and nonconscious mimicry' extending beyond a perception-behaviour link.

Emotional disposition and context

The impact of social context has been investigated by Bourgeois and Hess (2008), who used facial electromyography (EMG) to reveal participants exposed to different facial displays show

facial mimicry of sad expressions eliciting empathy but mainly from in-group members and facial expressions of anger was not mimicked. It is suggested only facial expressions serving an affiliative goal are imitated (Duffy and Chartrand, 2015).

An interactive artwork, called the Chameleon Project, uses facial recognition software to detect the emotional expression of the audience and displays video portraits of matching facial expressions from a database built by the artist. The facial expression of the audience triggered expressive portraits following an algorithm that avoids repetition in the display. These were based on probability rules that favour mimicry or complementary responses depending on the expression. During the 10-20 minute interactions the viewers expression was compared with the expression displayed in the portrait to determine whether a loop of facial mimicry emerged. The results showed a clear mimicry pattern for over 74% viewers in response to happy expressions. The effect was substantially less for sadness with less than 65% of viewers mimicking sad expressions. Other emotional displays elicited counter-expressions. Anger elicited mainly counter-expressions of mostly neutral reactions with some expressions of sadness and surprise. Disgust also elicited neutral reactions and surprise elicited happy reactions. This is again congruent with results fromBourgeois and Hess (2008) showing that for negative emotion the mimicry response is inhibited or can elicit a counter-expression. Exit interviews implied that an emotional bond and sense of intimacy ensued from interacting with the piece, or at the very least, the experience elicited self reflection on the audiences emotional state(Iacobini et al., 2009, 2010).

A study from Likowski et al. (2011) finds that facial mimicry of both positive and negative expressions is hindered by a sad mood. Other factors such as social anxiety hinders mimicry responses (Vrijsen et al., 2010) and feelings of guilt elicit more mimicry (Martin et al., 2010).

3.2.5 Mimicry distinctions

It should be noted that there is a disparity in the specific distinction of what constitutes as mimicry in the studies reviewed in this chapter. Although all studies observe matching nonverbal behaviour, the different types are evident in the list above. There is also an inconsistency on the time restriction placed on matching behaviours that are considered as mimicry in these studies. Studies observing behaviours such as foot-shaking or face-touching (self-adaptors)

are usually conducted by a group of researchers that propose the chameleon effect (see Section 3.2.2)as basis for mimicry, these place no time restriction on when a matching behaviour is defined as mimicry, counting the frequency of a particular behaviour over the duration of an interaction. Comparing the frequency of the particular behaviour in experiments were confederates performs the behaviour to control experiments where they do not. This means there could be an indeterminate amount of time between the stimulus behaviour from the confederate and the mimicry behaviour (restricted to the time taken to participate in the study). This has the limitation that it is not clear whether any instances of the measured behaviour that occur before the first stimulus behaviour by the confederate are counted in some studies. Other studies, particularly facial mimicry, head movement and postural mimicry define mimicry as matching behaviour that occurs between 0 to 10 seconds (depending on the study) of the mimicked behaviour.

Of studies on behavioural mimicry in adult interactions detailed above all analyse mimicry (or counter mimicry) in terms of it's prosocial effects, mimicry is either measured in a participant who either interact with a confederate exhibiting particular behaviours or observe photographs or videos of behaviours or facial expressions; a participant is mimicked by a confederate, or in a couple of studies by an avatar imitating the head movements of participants; or the participant is explicitly asked to mimic a confederate. In this set of research on the prosocial effect of mimicry very little studies analyse spontaneous natural interactions between naive subjects (for an exception see Griffin et al. (2015)).

3.2.6 Posture

3.2.6.1 Posture as indicative of social cues and affective states

Body orientation and posture have long been regarded as an indication of affective state, attitude and social status (Mehrabian, 1969).

For example, open body postures are related to positive affective states and attitude. Adopting an open postural stance indicates happiness (Kleinsmith and Bianchi-Berthouze, 2007; De Silva and Bianchi-Berthouze, 2004), honesty, open postures make the person speaking appear more influential and persuasive (Pease, 1981), whereas open postures make recipients appear more attentive (James, 1932; Pease, 1981; Lynn and Mynier, 1993). Openness of posture is said

to be regulated by status, adopt more open postures to persons of higher status (Mehrabian, 1969; Pease, 1981). Arms-akimbo is an exception as a negative open stance seen as aggressive, arrogant or proud. This posture is generally used with disliked recipients and recipients of a lower status (James, 1932; Mehrabian, 1969).

Contraction on the other hand is related to negative affect: withdrawn postures are generally associated with depression or repulsion (James, 1932), boredom (Beattie, 1988), sadness (Thrasher et al., 2011) and removing oneself (Scheflen, 1972). Arm barriers, such as arms crossed over the chest, are generally read as defensive to their recipient, however their performance is usually related to comfort (Beattie, 1988; Pease, 1981).

Bianchi-Berthouze et al. (2006) found pain was expressed by postural cues for a communicative effect, revealing how the pain is affecting the movements of the body or how it effects their affective or emotional state: 'it has been observed pain behaviour often increases in amplitude in the presence of solicitous others and/or health professionals', suggesting these expressions may not be directly caused by the pain itself, but more by the wish to share it with others.

In summary, body postures are seen to embody the affective state and attitude of their holder, regulated by status and relationship, giving cues for social understanding. It is noted postures adopted by persons of varying status', attitudes and cultural background would be expected to be different between different people. These regulative influences could see postures diverge- for example an aggressive posture is more likely to see an interactional partner adopt a complementary passive or defensive posture (Tiedens and Fragale, 2003)- this is incompatible with an automatic behavioural mimicry argument.

3.2.6.2 Posture congruence

Despite the above indicating posture is governed by factors with the potential to be asymmetrical, the persuasiveness of posture congruence has been widely noted (see in Section 3.2.4). Posture congruence has consistently been positively correlated with liking, rapport, shared viewpoints and agreement.

It could be argued posture congruence is a consequence of the perception-behaviour link mechanism as is the behavioural mimicry detailed in Section 3.2.1. Although studies on facial mimicry point to the notion adopting similar facial expressions facilitates an empathetic understanding

of the associated affective state, it remains to be seen if adopting the posture of a interactional partner can facilitate an understanding of the affective state their posture embodies. One study pointing to this, although minimal, finds adopting a posture taken during a previous experience or event enhances recall of that experience, it is argued posture and facial expression facilitate recall of context information if their motor configuration is congruent with the motor configuration adopted when encoded (Dijkstra et al., 2007). Posture congruence has also been shown to facilitate spatial perspective taking between interlocutors of differing spatial orientations (Urgesi et al., 2006). Both studies are in line with the simulative account - perception and action have overlapping, shared or associated neural representations.

Shockley et al. (2003) report the postural dynamics of one person can constrain the postural dynamics of another when in an interaction. Postural dynamics are seen to entrain to vocal patterns and breathing rhythms latent with a speech signal, this sees the postural sway of a pair converge when interacting. Even when interlocutors cannot see one another (Shockley et al., 2007). This suggests an alternative explanation to the ubiquity of posture congruence in interaction to the perception-behaviour link mechanism.

LaFrance (1979, 1982) proposes posture congruence has a communicative function, marking understanding and attention, suggesting interlocutors entrain postural similarity and co-ordination to signify comprehension- conversely a lack of congruence to signify inattentiveness or a potential problem in understanding. In this way, subtle posture adjustments serve as a structural feedback signal contributing to intersubjectivity between interactional pairs.

3.2.7 Gesture

Co-speech gestures are frequently used in conversation to complement and provide additional information to the accompanying speech (Kendon, 2004; McNeill, 1992). Temporally and semantically coupled with the verbal elements raised in speech, these embodied depictions can potentially provide a more direct representation of the imagistic and embodied aspects of a message (Goldin-Meadow, 2005). Gestures are frequently used to depict the visual or spatial domain, due to their own physicality, providing a more direct expression of an embodied experience through direct displays or demonstrations of aspects of an experience than speech (Gerwing and Allison, 2009).

The function of gesture within an interaction has been disputed. Two main approaches can be identified. The first approach is gestures are performed to aid speech production when speaking. Krauss (2000); Krauss and Chen (1996) have argued an important function of gesture is to facilitate lexical retrieval, suggesting the features represented in motor form by gesturing facilitates lexical access via 'cross-modal priming'. Taking evidence from findings that gesture inhibition reduces fluency, gesture still occurs whilst on the phone and lexical gestures usually precede their lexical affiliate (Butterworth and Beattie, 1978) they conclude gesture functions to aid articulation. de Ruiter (2000) point out gesture and its' lexical affiliate could be aligned because people switch modality as they anticipate lexical access problem to better convey their message. Furthermore, Beattie and Coughlan (1999) report counter-evidence to this hypothesis, finding there is no difference to lexical access in 'tip-of-the-tongue' scenarios when participants could gesture freely or had their arms crossed so gestures inhibited.

Alibali (2010) propose gesture facilitates conceptual planning. They base this hypothesis on a study that found participants focus more on perceptually present aspects of a narrative description when they could free gesture, than when their gestures were inhibited. They suggest this is because gestures are derived from simulated actions of content - perceptually present features become more concrete. Moreover, Gentilucci and Dalla Volta (2007) reports neurophysiological data favours the hypothesis speech and gesture share a common control system so can mutually influence one another delivery.

The second approach is supported by numerous experimental studies providing evidence for the communicative function of gesture (Bavelas et al., 1995, 2002). Kendon (2004) describes gestures as refining, qualifying, supplementing and giving a more restricted meaning to accompanying speech. Moreover, Holler and Beattie (2003) suggest gesture is used to disambiguate speech, observing more descriptive gestures when a person is asked to clarify meaning in their talk. Beattie and Coughlan (1999) also observe dialogue is more informative when gestures can be seen than not.

A gesture is deemed to be 'redundant' if it echoes information articulated in speech or does not contribute semantic information. Gerwing and Allison (2009) ran a redundancy analysis on conversations containing co-speech gestures, they observe in general, gestures are not redundant and do contribute information not present in speech. Bavelas (2007); de Ruiter (1995) observe gestures produced whilst interlocutors were on the phone or tape-recorder were redundant in

conveying additional information than the speech.

To measure the effect of visibility of gestures in an interaction Alibali (2001) asked participants to describe a previously watched cartoon to another participant whilst either sat where there was a screen blocking the visibility of their gesture, or without the screen. They found a lower frequency in representational gestures- gestures conveying semantic information of the message- when recipients could not see the person speaking, but beat gestures - gestures only conveying information about the production of the speech itself but not semantic information (redundant)- were not affected. This suggests certain types of gesture are specifically designed to be communicative to a recipient. Additionally, Özyürek (2000) found gestures are spatially orientated to their recipients.

Holler and Stevens (2007); Holler and Wilkin (2009) investigated the influence of common ground between interlocutors and gesture use in a narrative task. They hypothesise if a recipient is known to be aware of the content of a message then the current speaker will reduce their gesture frequency as there is less 'communicative demand'. They found a higher frequency of gestures describing size for conveying new information, however if a message is already known by a recipient there was less semantic information in speech but marginally more gestures produced which does not support their hypothesis. The latter finding is in support of the hand-in-hand hypothesis (So et al., 2009) - combining speech and gesture is less communicative effort. In a situation where the message is already known, gesture can serve as an abbreviation of the prior full articulation.

Jacobs and Garnham (2007) found an effect of communicative demand on gesture production in a narrative task, gesture frequency is higher for attentive over unattentive recipients, and higher for recipients for whom the message was new information.

3.2.7.1 Gestural mimicry and entrainment

Thus far, the accounts of behavioural mimicry analysed above have not considered behaviour conveying semantic meaning, such as representational co-speech gesture. The following section details evidence and examples of gestural mimicry in this context.

Tabensky (2001) highlights examples of gestural mimicry in live interactions between students. In these examples, aspects of the original gestures are recycled into a 'rephrasing' - it is sug-

gested these are reproduced in order to show understanding and agreement of the original message. Kimbara (2008) demonstrated gestural mimicry occurs in a co-narration task. Participants partook in joint narrations of a cartoon to a passive recipient. Co-narrators were either separated by a screen so they could not see their partner (but could still be heard), or both visible and audible to one another. Co-narrators performed more similar hand-shapes in the visible condition- it is suggested co-narrators align their decriptions to maintain coherence in a joint narrative.

Mol and Krahmer (2009); Mol et al. (2012) investigated whether the form of representational gestures are adapted to converge with previous gestural contributions from an interactional partner. Entrainment has been found in the verbal domain when linguistic forms are repeated across conversational partners, known as lexical entrainment or alignment. The repeated element takes the form of a referring expression and is mutually adopted to establish a common representation of meaning between interlocutors, forming a *conceptual pact* (Clark and Brennen, 1991) (although see Healey et al. (2010) who show interlocutors diverge on syntactic choice).

Initial studies involved participants watching a videotape of an actor describing the narrative of a cartoon, there were two versions of the tape, in one version the actor used co-speech gestures whereas in the other the actor did not use co-speech gestures. Participants were then asked to retell the narrative of the cartoon to the experimenter. Their results show more gestures where produced after watching the tape with co-speech gestures and these gestures correlated with the form of the gestures performed by the actor. Similarly, Parrill and Kimbara (2006) reports after viewing mimicked gestures in a video participants are more likely to reproduce these mimicked gestures in subsequent interactions. It is suggested priming may underlie these repetitions, rather than the construction of shared meaning across interlocutors.

Mol and Krahmer (2009); Mol et al. (2012) conduct a second experiment that also had two conditions, where participants watch a videotape of an actor describing the narrative of a cartoon. In the first condition the description was accompanied with congruent gestures (matching semantic meaning with speech), in the second condition there was only one gesture- which did not match the meaning of the concurrent speech. Participants repeated gesture forms more frequently if they had been presented to them in a linguistic context in which they were meaningful. Mol and Krahmer (2009); Mol et al. (2012) argue gestural mimicry is related to the

meaning they convey. McNeill (2005) describes this relation as the 'growth point':

> The smallest unit of the imagery-language dialectic is posited to be a 'growth point', so named because it is theoretically the initial unit of thinking for speaking out of which a dynamic process of organisation emerges. A growth point combines imagery and linguistic categorical content, and the theory is such a combination initiates cognitive events. A growth point is an empirically recoverable idea unit, inferred from speech-gesture synchrony and co-expressiveness. (McNeill, 2005).

A third experiment explored route direction gestures. A participant watched an actor describe a route using one of two conceptual ways of gestural depiction (perspective) but the verbal description was the same, participants were then asked to retell the route to the experimenter. Participants not only copied the actors' conceptual strategy but also the gesture form and the meaning of what each individual gesture referred to (Mol and Krahmer, 2009; Mol et al., 2012). These findings support the idea interlocutors converge on the certain gestural (and verbal) form they use to express these shared concepts, or growth points.

Kimbara (2006) terms this sort of gesture sharing contributes to 'form-to-meaning' mapping. Once a gestural form is associated with meaning, together they are added to the common ground- producing a shared gesture-speech unit. Holler and Wilkin (2011a) extend this by suggesting reusing one another's gestural forms facilitates a mutually shared understanding of referent expressions. In studies with pairs of participants describing symbols to one another in order to match them - mimicry was found in seven out of the 8 pairs. It was noted gestural mimicry was used to assert acceptance of referent and demonstrate shared understanding has been reached.

To test whether people attune their gestures to each other in a different interactional context, Furuyama (2000) used a different approach to the usual narrator-centred paradigm. Participants either took the role of an instructor or a learner in a dyadic instructional interaction task with origami (without the paper or any utensil). The hypothesis was learners would attune their gestures to those performed by the instructor as a learning strategy. Observations show firstly learners produce similar representational gestures to instructors and would elaborate them with extra elements. The instructor then takes on these additional elements. Secondly, learners synchronise their gestures to the instructor's meaningfully affiliated speech, gesturing

over the instructor's speech. Thirdly, learners mimic the instructor's gesture only when the focal side of the gesture faces the learner. This concludes gestural mimicry in this sense is a sharable material carrier of thought- sensitive to interpersonal factors and context. The apprenticeship setting positively influences gestural adaption and mimicry- as the purpose is to copy in order to learn (Furuyama, 2000).

In summary, the above literature suggests mimicry of gestures representational in form operate differently from the nonconscious behavioural mimicry said to be facilitated by the perception-behaviour link. Gestural mimicry functions to establish and share common gestural referents of meaning - used to clarify, establish and display understanding.

3.2.7.2 Motor mimicry as a strategic communicative mechanism

For the past 25 years Janet Bavelas and her colleagues have conducted research into what they term *motor mimicry*. Defined as the mimicry of an expressive behaviour, such as wincing when seeing another wincing in pain. Or the performance of an *expected* expressive behaviour in the perspective of another. For example, if one was to see another stub their toe on a door frame, one has the tendency to flinch one's foot and grimace in pain even though the event didn't not happen to oneself. Conceptualised as primitive empathy, Bavelas et al. (1986) describes it as an automatic reflex of conditioned cues based on own prior experience. Motor mimicry serves as an expression of the perceived emotion, an interpersonal act to put across 'I feel as you do'. These expressions are claimed as a form of nonverbal communication and not just manifest from an internal state. They hypothesise motor mimicry should appear more when the individual conveying the emotion is present and interacting.

Bavelas et al. (1986, 1987, 1988) ran a series of experiments to test the communicative function of motor mimicry. One of which recruited subjects under the guise of setting up an experiment, subjects watched the experimenter drop a heavy television set onto an already injured bandaged finger. At the point when the television was dropped; one condition saw the experimenters make eye contact with the subject; in the second condition the experimenter's eye gaze was directed away. Motor mimicry was more prominent when the experimenters made eye contact at the subject.

The presence and availability of a receiver affected both the pattern and timing of

> motor mimicry. As the probability of eye contact with the receiver increased, motor mimicry not only increased generally but was available at the best 'delivery point.' When, on the other hand, it became less and less probable the victim could see the expression, motor mimicry ether faded away quickly or did not occur at all. These observers' faces seemed to go on hold, apparently waiting for eye contact that never happened (Bavelas et al., 1986, p. 325).

It is noted motor mimicry occurs, albeit not at full strength, without the presence of a recipient. For example in response to stimuli from a videotape. This is not regarded to be a byproduct of a mirror system as mentioned above, but the authors suggest these responses form a noncommunicative base rate of motor mimicry, and is analogous with 'talking to oneself'. Suggesting as the response is skilfully and precisely inserted into an interactive sequence, it is unlikely to be the result of an overflow from an inner mirror mechanism. Motor mimicry as a nonverbal communicative is the simplest way to establish intersubjectivity, requiring only bodily expressions to demonstrate reciprocity and mutual understanding (Bavelas et al., 1986, 1987, 1988; Bavelas, 2007).

3.2.8 Feedback behaviours

The occurrence of recipient feedback is thought to facilitate the incremental process of a conversation as a joint activity (Heylen et al., 2011; Ward and Tsukahara, 2000; Allwood et al., 2007; Allwood, 2003; Cerrato, 2002).

Feedback is essential for intersubjectivity as it provides signals demonstrating the success of an interaction as it unfolds (Clark, 1996). Communication is regulated by interpersonal feedback, the control is shared between two (or more) individuals as a collaboration. Each expression is directly predicated moment-to-moment on the previous action. Often thought of as a fluency index- feedback helps a speaker formulate their speech in accordance to how a recipient is responding to their talk; allowing speakers to judge how their talk is being received as they speak (Cerrato, 2004). Bavelas et al. (2000, 2006) identified the level of feedback produced by a recipient directly influenced the fluency and effectiveness of the speaker's turn. Storytellers communicating to distracted recipients speak less fluently and stories are less captivating than those whose recipients are attentive. Stubbe (1997) observes how feedback signals are used and

perceived by different cultural groups. Differences in how groups interpret shared understanding indicate the highly collaborative and systematic nature of recipient feedback.

Previous research in this area identifies several standard categories of recipient feedback serving incremental functions: *Contact, Perception, Comprehension and Attitudinal/Emotional* (Allwood et al., 1993; Cerrato, 2004, 2007).

Contact/Perception:

Contact and perception feedback show a continuation of recipient attention acknowledges there is a message being put across. For example, generic nodding or vocalisations such as 'yeah' or 'mmhmm' do not interrupt to take the floor but can sometimes seek continuation as a way of avoiding the floor. Cerrato (2002) classifies this type of feedback as a subtype of back-channel feedback expressions, otherwise known as continuers. These share the following features:

- responds directly to the content of another's utterance
- are optional
- does not require acknowledgement

This definition rules out: post completion vocalisations, feedback occuring just after an utterance that could be from 'reflecting on some cogitation', and the answers to questions (Cerrato, 2002).

A number of studies have suggested back-channels are closely aligned to a speaker's talk. Ward (1996) found prosodic cues in speaker's utterances can predict the production of recipient back-channel feedback in Japanese dialogues. These are mainly related to syntactic features in the speaker's utterance such as at the end of a grammatical clause, indicating each unit has been understood so the speaker knows to move on to the next. Rauzy et al. (2007) extended on this by analysing what interactional events influence back-channel production in a corpus of French storytelling dialogues. Prosodic cues typically occurring the end of a grammatical clause are also followed closely by recipient back-channels in this French corpus. Back-channels also closely follow intonation that typically elicited a validation of new information but also considered a turn-holding cue. Gestural back-channels, such as nodding, were more likely to occur when the person speaking is looking at their recipient (Goodwin, 1979; Bavelas et al., 2006; Rauzy et al., 2007). In an interaction with more than one recipient- the primary recipient (the interlocutor

the speaker is currently looking at) nods more than secondary recipients (other interlocutors in the conversational party) (Healey et al., 2013). This suggests feedback of this type is highly functional and are directed specifically to aid the speaker's delivery.

Comprehension:

Another function of feedback is to acknowledges understanding of a message. Comprehension feedback is sometimes difficult to distinguish from contact and perception feedback. The clearest cases are when the feedback is in the form of a direct referent to their understanding, for example 'I see', 'Aaaah', 'Oh right'.

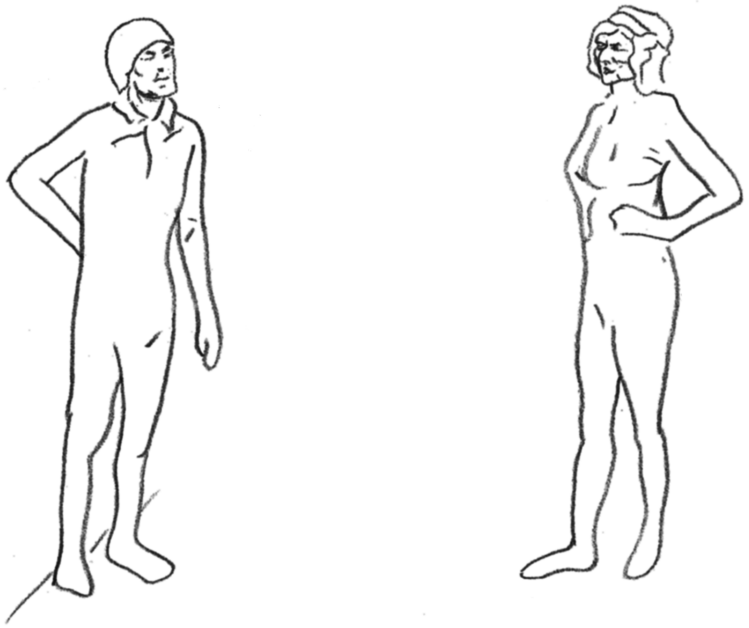
Attitudinal or Emotional:

Another form of feedback is attitudinal or emotional, expressing a point of view or attitude towards the message. Schroder, Heylen and Poggi (2006) identified the subtype of recipient responses displaying attitudinal or emotional feedback called affect bursts. Affect bursts are 'brief, discrete, nonverbal expressions of affect in both face and voice as triggered by clearly identifiable events' (Schröder, 2003). Their experiments collecting recognition ratings of vocalisations of such phenomena indicated affect bursts serve to display emotions gratifying for the speaker, or show empathy toward the speaker but generally never expressing a negative attitude or emotion toward the speaker (Heylen, 2006). Similarly, Bavelas et al. (1987) classify empathetic recipient responses as motor mimicry (see Section 3.2.7.2).

The following example shows an emotional or attitudinal feedback. The participant on the left, labelled CH in Figure 3.1 is describing an event in lines 1 to 7 of something dropping a heavy object from a stairwell and it falling onto his back in between two vertebrates in his spine- 'so he dropped a er I guess about one and a half to two kilo piece of er galvanised steel onto my back from the err upstairs down a stairwell and so that dug in between two vertebrates'. In response the participant on the right, labelled NCH cries 'What?!' in line 5 and groans in pain in line 8 '.hhhhh', she's quickly slaps her hands to her cheeks in surprise on the word 'What?!' and they remain there for the rest of the sequence. Her facial expression is in a surprised state, eyes wide, mouth and lips in the shape of an 'O'. This demonstrates shock to his descriptions, and well as discomfort when she groans in line 8 '.hhhh', both dramatically expressing an emotional and attitudinal response to his utterance in the form of an 'affect burst'. She does a version of what Wilkinson and Kitzinger (2014) terms 'ritualised disbelief' - these expressions do not signify

'genuine' surprise, but demonstrate an appropriate response - she is not reliving the surprise for herself, but using its display as a communicative resource to indicate she understands the severity of the event. When she groans she is demonstrating she understands the pain the event would have caused, groaning the way the participant might when the event actually occurred, placing her perspective as his for the sake of demonstrating his pain. This can be viewed as an instance of Bavelas' motor mimicry as the expression functions as feedback within the interaction and not because she is under the same discomfort the event might have cause, nor is she surprised to the point of panic, but she enacts these expressions to demonstrate her understanding of the event her interactional partner describes. This goes beyond indicating contact, perception and even comprehension as the action demonstrates she sympathises with her partner.

Figure 3.1. Wincing sketch



1. CH: so he dropped a er
   (1.3)
2. CH: I guess about one and a half to two kilo piece of er galvanised steel
   (0.3)
3. CH: onto my back
   (.)
4. CH: from the err=
5. NCH: =What?!
   (.)
6. CH: upstairs down a stairwell
   (0.5)
7. CH: and so that dug in between two vertebrates
8. NCH: [.hhhhh]

Both affect bursts and motor mimicry contain emotional or attitudinal responses that occur simultaneously to the speaker's utterance. It would be expected that descriptions of sensory experience would provoke empathetic responses like motor mimicry, especially during descriptions of pain. Moreover, empathetic responses should be most likely to occur when the recipient has a good understanding of the sensation.

3.3 Conclusion

As reviewed in this chapter, numerous studies have found evidence of the ubiquity of behavioural mimicry in interaction. Studies have also demonstrated marked prosocial effects relating to behavioural mimicry within interaction, such as empathy and rapport. It is claimed that a tendency to mimic one another is an automatic response to perceiving the actions of others, termed the perception-behaviour link (Chartrand and Lakin, 2013).

The perception-behaviour link draws upon cognitive theories as the basis of how it functions, suggesting that perceived actions and executed actions share a common representational cognitive domain or overlapping cognitive semantic 'schemas' (Prinz, 1990; Barresi and Moore, 1996). The mirror neuron system is hypothesised to provide neuro-physical evidence for the perception-behaviour link, neuroimaging studies demonstrate overlapping sensory and motor representational codes when a subject both perceives and executes an action (Iacoboni et al., 1999; Decety and Chaminade, 2003b; Decety and Sommerville, 2003; Wang and Hamilton, 2012). It is proposed that behavioural mimicry occurs as representations of action automatically lead to actual behaviour after a certain threshold (Chartrand and van Baaren, 2009).

Gallese (2013) extends the function of the mirror neuron system to playing a role in understanding the affective states of Others in a theory of 'embodied simulation'. As representations of action and affect are bodily in format; the mirror neuron system allows one to reuse their own mental states or processes as a vehicle to functionally attribute mental states to others. It is this functionality that is said to underpin empathy and intersubjective understanding.

The hypothesis that unites the perception-behaviour link, the mirror neuron system and observations of behavioural mimicry in social behaviour, suggests that the inward simulation of actions and expressions of Others is key to an intersubjective understanding; it is the mech-

anisms behind this simulation that account for the propensity to engage in nonconscious and automatic mimicry.

Part II

Methods

Chapter 4

General methods: Analysing multimodal interactions

4.1 Aims of analysis

Communicating the character of a felt sensation such as a pain seems harder than communicating about, say, the concrete events that caused a pain. Felt experience can be difficult to articulate and there is no guarantee that it is shared from person to person. Nonetheless, we appear to have the capacity to empathise with each other's experience, although the particular mechanisms behind empathy are still disputed and unclear (Preston and de Waal, 2002). Much of the work reviewed in preceding chapters concerns the in principle possibility (or impossibility) of knowing another's experience. This work attends to the empirical question of: how people express their experiences to each other; what recipients do when they try to understand and subsequently demonstrate an understanding of another's experience in conversation; and how this process is collaborative. In particular, how we employ our embodied resources such as gestures, postures and expressions.

The purpose of this study is to examine what mechanisms people use to come to intersubjectivity. Whether what recipients do is consistent with embodied simulation; testing the theory that people automatically mimic or entrain to one another in order to understand and demonstrate their understanding of how each other feels; is intersubjectivity reached via strategic communicative mechanisms? To discover how interlocutors use their embodied resources to

communicate two corpora of dyadic interactions in a laboratory setting are analysed.

The basic premise of this approach is that it is the use of our embodied resources in a live conversational exchange that is key to understanding how experience is communicated in practice, rather than telling a story to a camera or using an experimental confederate (Bavelas, 2011; Kuhlen and Brennan, 2012).

4.2 Corpus 1: Conversations about bodily experience

A corpus of natural, non-scripted interactions was collected, focussed on specific embodied experiences like the experience of laughing or the sensation of a headache. Following Bavelas' close-call story technique (Bavelas, 2011) participants are asked to describe to each other recalled occasions of the specific experiences, emphasising the embodied element within these memories. The target experiences are explicitly chosen to involve personal embodied experiences with either a negative (painful) or positive (pleasurable) valence. The inclusion both positive and negative experiences to provide a factor of valence after the suggestion that empathy should be stronger with more negative experiences. This can occur as a result of moving one more strongly towards sympathy and altruism (Preston and de Waal, 2002; Goubert et al., 2005) In all cases audio, video and body movement data is collected (see Section 4.2.4).

4.2.1 Pilot studies

Pilot studies were carried out to test the video and motion capture equipment, establish the best procedure for syncing the video to the motion capture data, the best video and motion capture camera angles, and the optimal set-up parameters for the equipment (see Section 4.4 for a description of the motion capture system) and to test the appropriateness of the instructions given to the participants[1]. The aim was to encourage participants to describe their experiences in a naturalistic way. One strategy of a game of 'guess the experience' was rejected; firstly as it made the participants nervous, and secondly because descriptions were truncated as soon as a correct guess was made. Simply asking participants to talk about particular experiences was the most successful instruction. To tune the task a range of alternative experiences were

[1]Six pairs were recruited to test these approaches/experiences

also piloted including heartburn, earache, a back click, thirst quenching drink, a good stretch, a dead leg, stubbed toe, a paper cut, a twisted ankle, brushing hair, having something stuck in-between teeth, pulling off a plaster. The experiences judged most commonly experienced were selected:

- a headache
- the taste of a nice meal
- a toothache
- a satisfying yawn
- a backache
- a back massage
- a stomach ache
- laughing out loud

The final shape of the task resulted in participants alternating to describe the 8 randomly ordered experiences to one another.

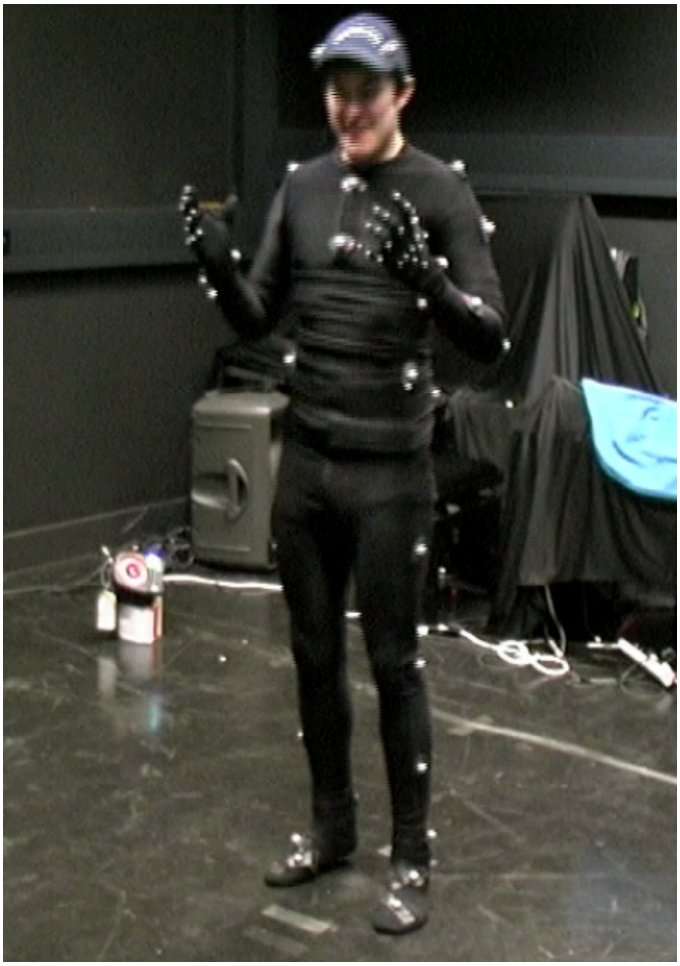
4.2.2 Participants: Corpus 1

A total of 24 participants were recruited. The recruitment process targeted undergraduate students of various subjects at QMUL, as well as on-line recruitment lists such as the Arts Jobs List and Gumtree. Participants ages ranged from 18 to 60, consisting of 12 females and 12 males placed in 12 pairs. The aim was to elicit unrehearsed and spontaneous descriptions in an interaction. Of the dyads, approximately half of them signed up as a friendship pair and the other half were randomly paired by the experimenter, so could be considered as strangers. Three of the pairs were both male, two were both female, and the remaining five were mixed.

4.2.3 Materials: Corpus 1

The corpus of speech, video and body movement data was captured in the Performance Laboratory at QMUL. Video footage included a full body, face-on view of each participant for the

Figure 4.1. In the lab, the participants wear a motion capture suit, cap and gloves



duration of the study (see Figure 4.3 for a diagram of the layout of the room). Motion capture data was also obtained for participants using a Vicon optical marker system. A set of cards were placed on a small table next to where the participants stood. Each participant was given a stack of these cards and asked to take turns selecting one card at a time. There were 8 cards in total per participant, each card had one of the selected experiences written on it (See Figure 4.2).

Figure 4.2. Example cards

4.2.4 Procedure: Corpus 1

Participants were given written instructions outlining the study procedure (See Appendix for participant information sheet). They were asked 'to recall some experiences and talk about them to each other.' The experiences were written on sets of cards and the participants alternated in describing the experiences on the cards to each other. When it was their turn each participant was instructed to do the following:

1. Take the top card from your set, read it and discard the card.
2. Try to remember a time you had the experience written on the card and recall the details of the particular sensations you felt at the time.
3. Explain the details of these sensations to your partner so they can understand how this experience felt to you.

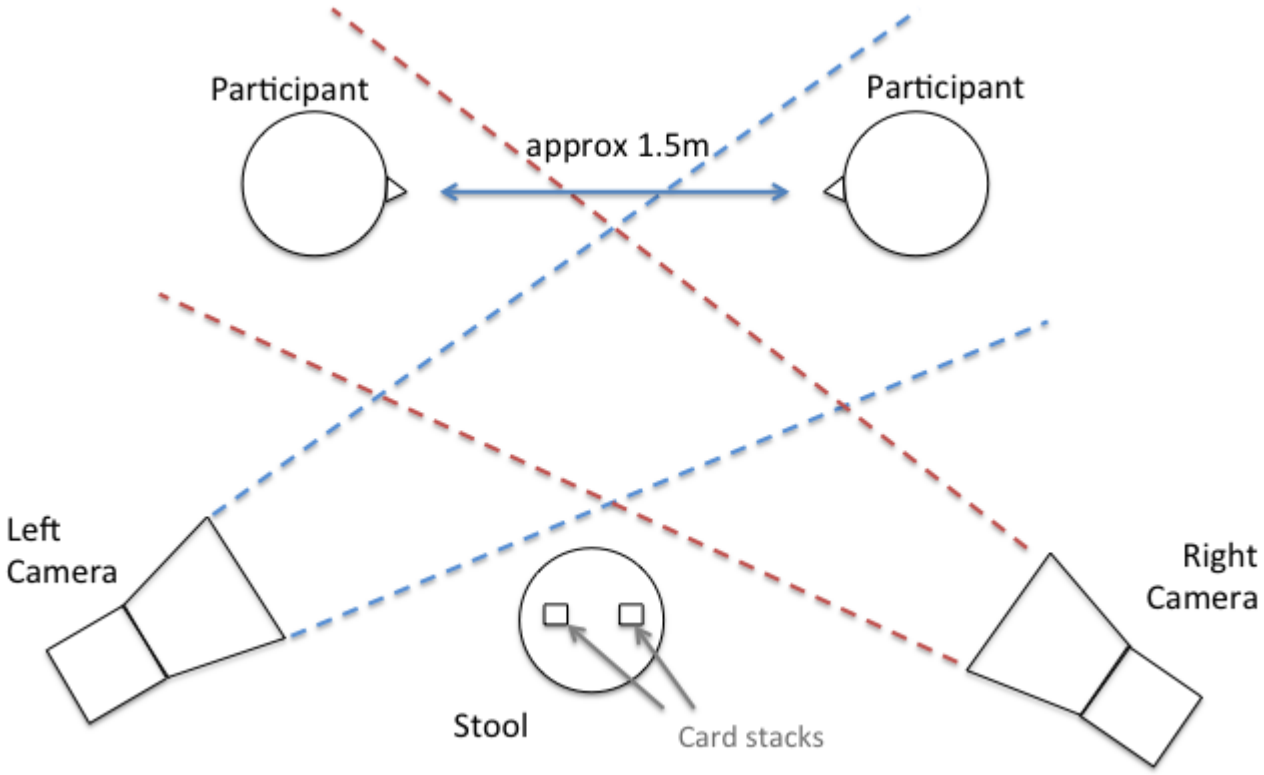
Participants were instructed to talk about their experience with their partner for no longer than 2 minutes. An emphasis was placed on describing how this experience felt or the particular sensation they felt at the time of the experience. On each description the listening participant was encouraged to talk and ask questions at any time, the process was described in the instructions

as an *exchange*.

As the study required participants to wear motion capture suits with reflective markers; there may be an effect of this on the interaction, such as participants recognition that their movements were being studied or merely an increase in self-consciousness. To mitigate this effect and allow participants to practise and settle into the irregular nature of having a conversation in the conditions of the lab, the first two experiences that came up in the set of cards (which were the same for every pair) in each session were practise trials that were not analysed (these were a headache and the taste of a nice meal). Aside from these two dummy experiences at the top of the stack, the cards were shuffled into random order for both participants in each session.

Participants were restricted to standing where their bodies and hands could be seen by both video and motion capture cameras.

Figure 4.3. Layout of lab



4.3 Corpus 2: Conversations about current affairs

A possible criticism of the 'body experiences' corpus is the peculiar nature of the task. Discussing sensations was regarded by some as a relatively unnatural form of conversation. Although the participants do talk about each other's experiences; the task, and the asymmetry of the storyteller/recipient roles could all influence their behaviour. To provide a partial check on these sources of bias a second corpus of dyadic interactions was analysed in order to compare results. This corpus, collected by Mary Lavelle (2010) is a part of a larger corpus of triadic and dyadic interactions - only the dyadic informal interactions were used. These involved two participants discussing their thoughts on current popular topics. Although this corpus still featured interactions inside a lab and participants still wore lycra motion capture suits in order to collect movement data the task was designed to emulate 'pub talk' with the intention of providing a relatively naturalistic form of conversation. Another significant difference between this corpus and corpus 1 is that the participants are sat on stools, potentially contributing to changes in forms of gesture and posture.

4.3.1 Participants: Corpus 2

A total of 20 participants were recruited. Participants ages ranged from 18 to 65, consisting of 7 females and 13 males placed in 10 pairs, three were all male and seven were mixed. Participants were recruited in a similar way to the previous corpus. Participants who were not fluent English speakers were excluded. Participants were unfamiliar with their interacting partners and had not met prior to the study. However, as this particular conversation was part of a larger corpus it occurred at the end of a longer session, so the participants had been talking to each other for around 30 minutes prior to these conversations, providing additional time to adjust to conditions in the lab. Previously in the session the two participants were involved in a debate about a hypothetical situation with one other participant. Although not instructed by the experimenter, some overflow of the structure of this prior task also occurred.

4.3.2 Materials: Corpus 2

Audio, video and body movement data was captured in the Augmented Human Interaction (AHI) Laboratory at QMUL[2].

Video footage was captured from two spots in the lab capturing each participant for the duration of the study. These included a camera in a top corner of the room, looking down on the participants from the side, and a side view of both participants at eye level. Motion capture data was also obtained for each participant using the same Vicon optical marker system (described in more detail in Section 4.4). Marker placement for both corpora was based on the same skeleton template in both corpora. Because they were seated participants only wore the top half of the lycra suit and participants leg movements were not captured[3].

4.3.3 Procedure: Corpus 2

A topic was provided at the start by the experimenter asking participants to discuss their opinion freely for 5 minutes. These topics included discussing the impact of the London Olympics on the city (of which they were residents) and a popular reality television programmes they knew of or had seen. Participants were seated on stools (See Figure 4.4).

[2]The AHI Lab is much smaller than the Performance Lab used in Corpus 1. While a larger room is beneficial for motion capture data as the area the cameras cover is bigger in the larger room, it is inconsequential in this case as the capture space used in both corpora was a similar sized area that the participants occupy, resulting in no difference to the quality or type of motion capture data collected.

[3]This difference does not effect the quality or type of motion capture data collected, however there will be no point of comparison between the corpora for the leg movement data collected for the participants in Corpus 1.

4.4 Motion capture as a tool for analysing embodied expressions

Body movement data was collected using a motion capture system (see Section 4.4), which has the advantage of providing precise three dimensional measurements of each participant's movements accurate to the millimetre, and fine-grained enough to provide data points at up to 120 frames per second.

The system used for both corpora is an array of 12 Vicon MX infrared or near infrared cameras, driven by Vicon iQ software. This particular equipment uses a passive tracking technique which frees participants from the restriction of tethered to wires used in active systems. Infrared reflective markers (rubber balls covered in high visibility tape) are attached to participants clothing. Each camera strobes infrared light onto the capture space to produce a 2-dimensional image of the reflections of each marker worn by the participant, at 50 frames per second. The set up requires ensuring that every marker can be seen by at least 3 cameras at any one time in order to facilitate triangulation of its position in 3-dimensional space. This produces 3-dimensional coordinates for each marker when the 2-dimensional infrared images are reconstructed in post-production.

4.4.0.1 Calibrating the system

Once the cameras are in position, there are three tasks that the Vicon iQ software must perform to calibrate the cameras. Firstly when the capture space is empty, the Vicon iQ software is set to the record the background and exclude any unwanted reflections in the room by setting a threshold. Secondly, so that the Vicon iQ software can determine where each camera is relative to each other, a pre-defined object is moved around the space until each camera has been triangulated. This object has at least 3 markers placed on it and software is programmed to recognise its function as a calibration tool. Lastly, placing a pre-defined object (a floor plate) on the floor in order for the software to determine the axes, scene origin and where the floor plane lies within the capture space.

Figure 4.4. Corpus 2: Sketch courtesy of Daniel Adderley

1. P1: Britain's got talent, [I don't watch X factor]
   P1: [(points to himself)]
2. P2: You don't watch it?
   (0.2)
3. P1: I don't watch X factor.

Figure 4.5. Motion capture optical markers

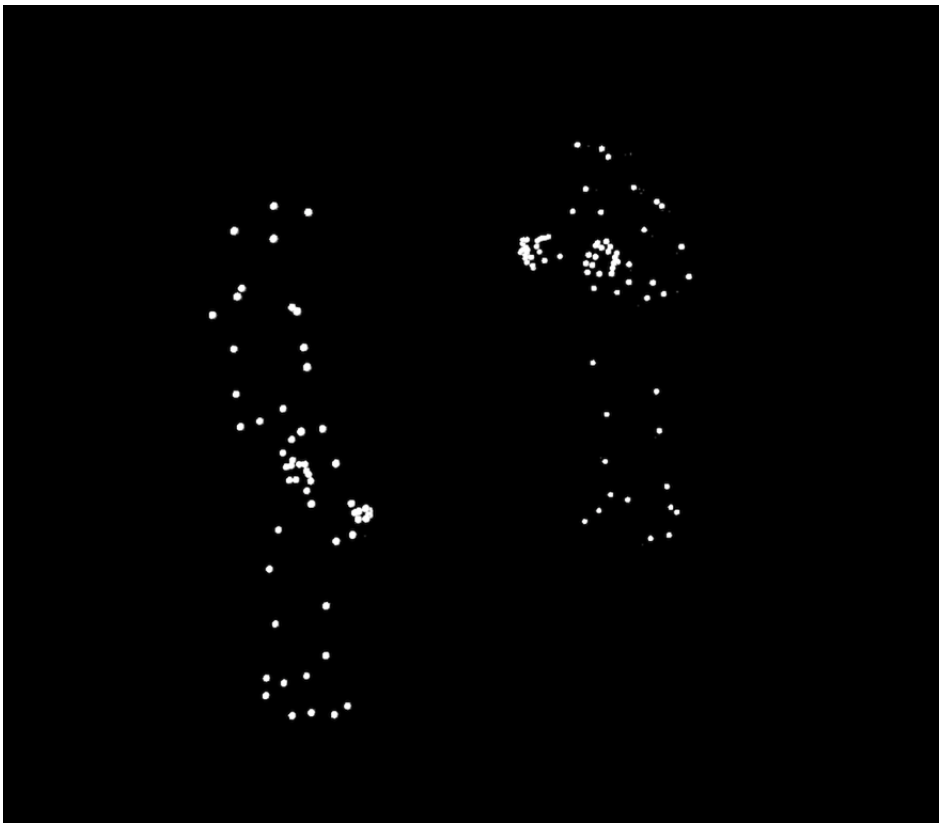


Figure 4.6. Wireframe skeletons

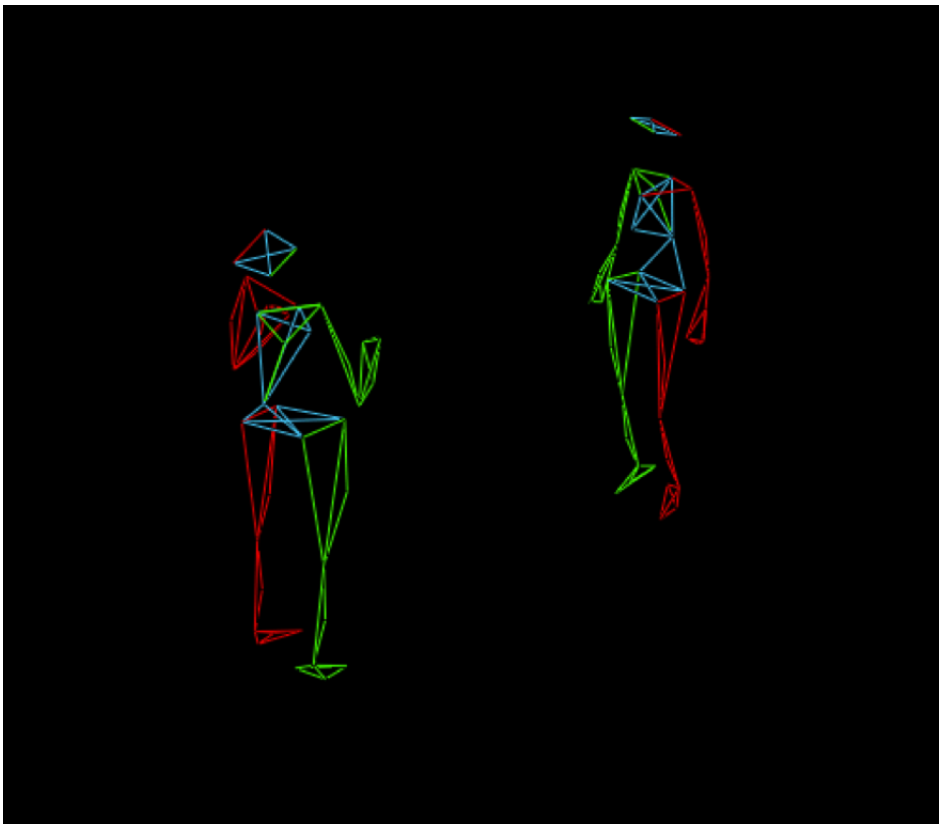


Figure 4.7. Front marker placement

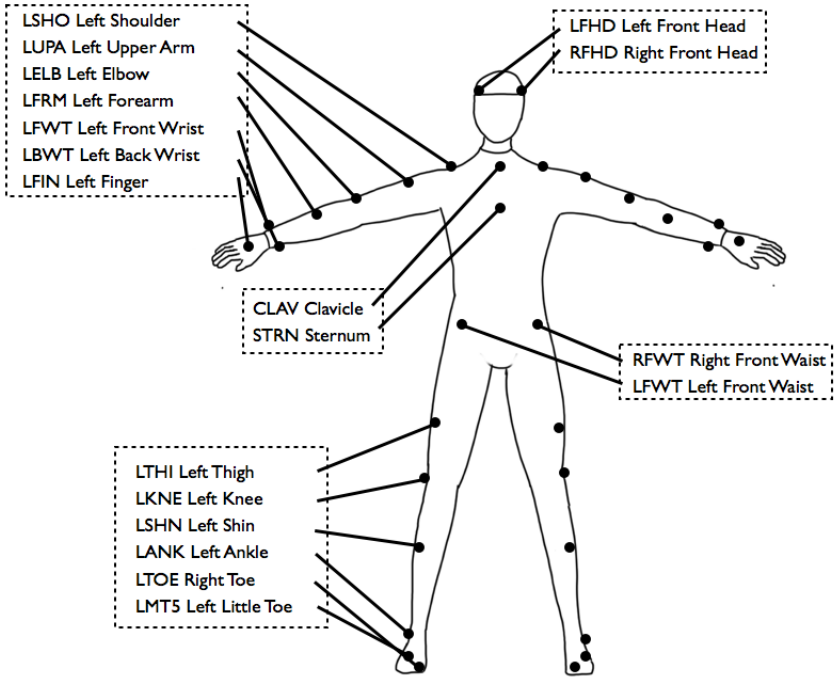
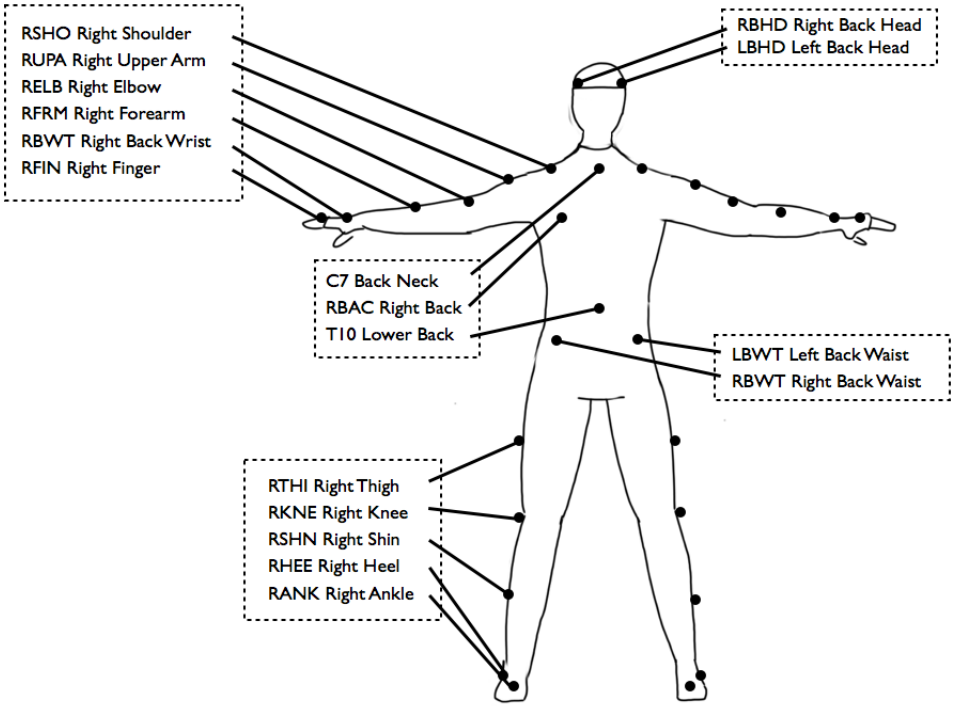


Figure 4.8. Back marker placement



4.4.0.2 Marker placement and VSK files

Each participant wore optical markers placed in accordance with the template illustrated in Figure 4.7 and Figure 4.8. These were secured to the lycra suit using velcro. Once the markers were in place a short captures of each participant moving to the limit of their physical range was recorded; to provide range of motion (ROM) captures. Vicon can produce a VSK file containing an individual skeleton template for each participant using the ROM capture, containing data about rigid bodies measurements and marker position restrictions due to each person's kinematic limitations (the limit to which one can bend). The VSK file is used during the processes of reconstructing and automatically labelling markers as described below.

4.4.0.3 Post-processing: Reconstructing and labelling movement data

After the capture is complete, the data starts off as a set of identical optical markers moving in 3-dimensional space (see Figure 4.5). To reconstruct the optical data into meaningful positional values each optical point is labelled in accordance with the skeleton template (see Figure 4.7 and Figure 4.8). The labelling process matches optical points to markers on the body, for example left upper arm or left shin. Once each marker is labelled a wire-frame representation of each participant's skeleton can be constructed (see Figure 4.6).

4.4.0.4 Processing and cleaning the motion capture data

Markers can be subject to drop-outs when a marker isn't in clear line of sight of at least three cameras and the markers position cannot be triangulated. Temporary occlusion of markers is unavoidable when capturing natural social interaction as movements are not restricted. Movements such as crossing arms in front of the torso, or standing holding arms behind the back occlude markers on the lower arms and hands from the cameras field of view. The Vicon iQ software provides a pipeline with three algorithmic processes to fill these gaps:

1. Splines: Reconstructed points calculated by predicting the trajectory of a point from it's position before and after the drop out. This is only used to predict the position of markers where the gap is below 10 milliseconds.

2. Rigid Body: Missing point are calculated where they have a fixed position in relation to

a group of other points.

3. Kinematic Fit: Calculates points incrementally in accordance kinematic constraints of the human skeleton template.

As a result of these computational processes the data is subject to a small degree of noise, caused by the interpolated gaps.

4.4.0.5 Exporting 3-dimensional data

Time series data with the position of each labelled marker and joint rotations in three dimensions are exported. There are two different file types exported from the software for analysis, one to analyse positional data and the other to analyse joint rotations.

Positional data is exported in a .trc file, a tab-delimited text file format. This contains positional values in three dimensional space in millimetres relative to the origin of the capture space. Static data such as frame rate, frame count and file origin are recorded in the header. Dynamic data is presented in tabular form. Each marker's three dimensional coordinates are listed across the rows; each marker occupying 3 rows (x,y,z values). For one participant a full skeleton of 43 markers occupies 129 rows. Giving 258 rows for both participants. Frames are represented in columns, resulting in a matrix of marker coordinates to frames numbers.

Joint rotation is exported using a Vicon specific file format (.V file). V files use a hierarchical structure in order to ascertain local orientations for each joint. Once again, static data such as frame rate, frame count and file origin were recorded in the header. The dynamic data section was represented as DOF (degrees of freedom) tags for each entity on each frame. Each DOF tag contains a record of the frame number, participant ID, joint entity, local x,y,z angle-axis values.

4.5 Post-processing motion capture data

4.5.1 Calculating a nodding index

Nodding is a very common feature of interaction which interlocutors produce feedback (see Section 3.2.8). A Python script[4] for detecting nodding-like movement developed by Battersby et al. (2010) provides an index of head movement designed to detect the characteristics of nodding. Direction changes in the vertical orientations can be mapped into the characteristic peaks and troughs of a nodding movement. The script detects a rate of change that is greater than 0.3mm per frame as a potential nod. Head movement signals such as prosodic body movement like shifting posture or swaying are filtered out by using a high-pass filter at (1Hz) before the nodding measure was calculated. In addition, sudden body shakes or high frequency movements caused by camera error are filtered out using a low-pass filter (4Hz). Resulting in a signal range between 1Hz and 4Hz (deemed ordinary head movements by previous research- for more detail see Battersby (2010)). This range is used for peaks and troughs resulting in a binary value for each frame.

4.5.2 Calculating hand movements

Another script from Battersby (2010) indexes hand movement by calculating hand speed from one marker distance in 3-dimensional space over time (mm/frame). To do this a threshold of the fastest hand over one standard deviation above the mean acceleration for the entire interaction is treated as an index of 'significant hand movement'. Informal comparison of this index to the corresponding video shows that the index usually only captures the 'preparation' and 'retraction' stage of a gesture, the result of raising ones hands into gesture position then lowering them, although these are still useful. Gesture are often held for short periods around the 'stroke' phase of a gesture to align meaningful content[5] with the corresponding speech (for example holding pointing gestures) and these are not captured by the index. It is good at detecting beats or baton gestures; gestures used for emphasis of certain words, related to the delivery of the speech rather than the content of the speech (McNeill, 1985; Efron, 1941). This

[4] https://www.python.org/

[5] Usually gesture does not occur during pauses in speech and the stroke phase of a gesture rarely occurs before the relevant point within an utterance (Butterworth and Beattie, 1978)

gesture type is not covered in the hand coded gesture taxonomy (See Section 4.6.5.2).

4.5.3 Providing a similarity measure for posture congruence

The motion capture skeletons make it possible to directly, qualitatively test the degree of posture mimicry found in dialogue. Yang et al. (2010) developed a simple method for comparing postures in motion capture data. This method produces a similarity measure for the angle between corresponding joints, for example the angle of my elbow joint is compared to the angle of your elbow joint. This value will fall between -1 and 1; -1 being fully dissimilar and 1 being identical. This technique uses local rotation matrixes for each joint, which means the angles are relative to their parent joint. For example, if I bend my elbow joint the angle will correspond in relation to it's parent joint, the shoulder joint. This makes the comparison more accurate across multiple bodies than absolute rotations. Relative angles do not depend on the direction people are facing and also normalises for body size. The complication of comparing relative joint angles is that this requires the data to have a hierarchical joint structure, so the V file format is used to perform this analysis (see Section 4.4.0.5).

4.5.3.1 Factors constraining posture

Physiological constraints

Shared human physiology entails similarity values tend to be high as the limitations of human body plan means it would be difficult for two people to adopt entirely dissimilar postures. For example, elbows do not normally extend or rotate more than 180 degrees. People have a basic body plan in common and physiological constraints on their range of possible movements- for this reason it is not expected for posture similarity to be zero between any two people.

Task constraints

Yang et al. (2010) tested and verified the success of this method from findings derived when comparing different yoga poses. They suggest it reasonable to regard two yoga postures as similar to each other, when the similarity score is above 0.8, considering the substantial differences in joint angles when adopting different yoga poses. When two people are engaged in conversation, the range of joint angles adopted during the interaction doesn't approach the

variability in joint angles produced for yoga poses - interactional partners very rarely stretch to the extremes of their movement when involved in talk, unless of course talk involves stretching to the limits of the body, which is not the task in either corpus. For this reason it is expected similarity values based on comparing two people engaged in conversation are likely to be higher than a 0.8 threshold found reasonable to consider yoga postures as similar. This demonstrates that situational factors can also inflate the degree of similarity observed. In the current task people are standing, so a reduction in the range of leg angles is to be expected as they are commonly constrained by the need to keep their legs in a position capable of supporting their body weight.

Interactional constraints

The format of the bodily experience corpus task structure ensured that one interlocutor adopts a speaking role- describing a recalled bodily experience. Their conversational partner would take the recipient role- the audience to the description expected to interact with the cardholder to ensure comprehension. It has been shown that speakers and recipients are distinct in their nonverbal behaviours. Speakers are more likely to produce gestures and expressions. Listening passively to the speakers' description ensues that recipients avoid gesticulating, producing little or no gesture during any backchannel feedback (Gullberg, 1998; Battersby, 2011). Effects like role will place interactional constraints on the posture adopted by either interlocutor.

Baseline reference values

In order to account for the relative contribution of physiological constraints, task constraints and interactional constraints on posture similarity three reference values are formulated:

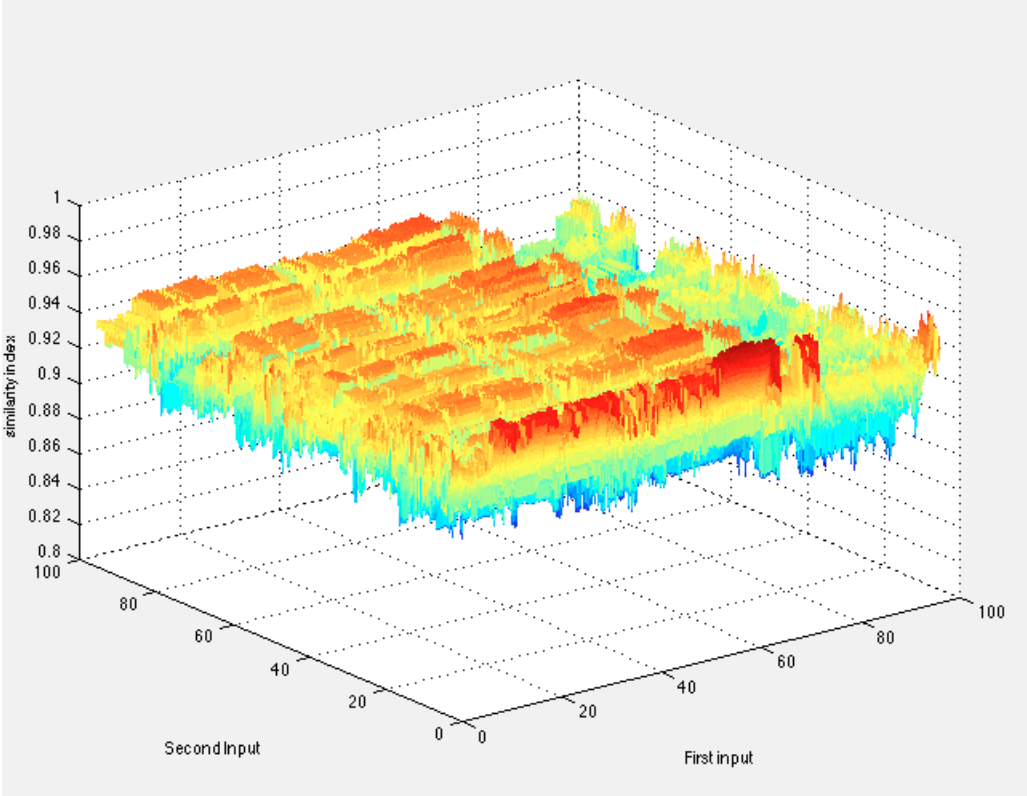
1. **Self-Identity**: Cardholder and recipient posture is compared with their own in the same interaction: producing a self-similarity value. Posture self-similarity is calculated for all offsets over the course of a single interaction. At a zero offset a cardholder/recipient will have a self-similarity of 1. At larger offsets idiosyncratic physiological constraints will always impact one own posture to increase self-similarity. A person's customary mannerisms or repeated characteristic postures or gestures will also increase self-similarity.

2. **Self-Min**: Cardholder and recipient posture is compared to themselves for all offsets in an interaction of the same topic but holding the opposite role. Producing a 'minimum' estimate of the influence of physical body constraints and characteristic movements on

posture across different interactions.

3. **Chance**: Posture similarity is calculated for all offsets for randomly combined pairs of cardholder and recipients (who did not interact with each other) the same task and the same topic. This provides an estimate of the 'lower bound' on posture similarity. Demonstrating the level of postural similarity due to the influence of a shared physiology or due to situational influences of the task- such as sitting or standing.

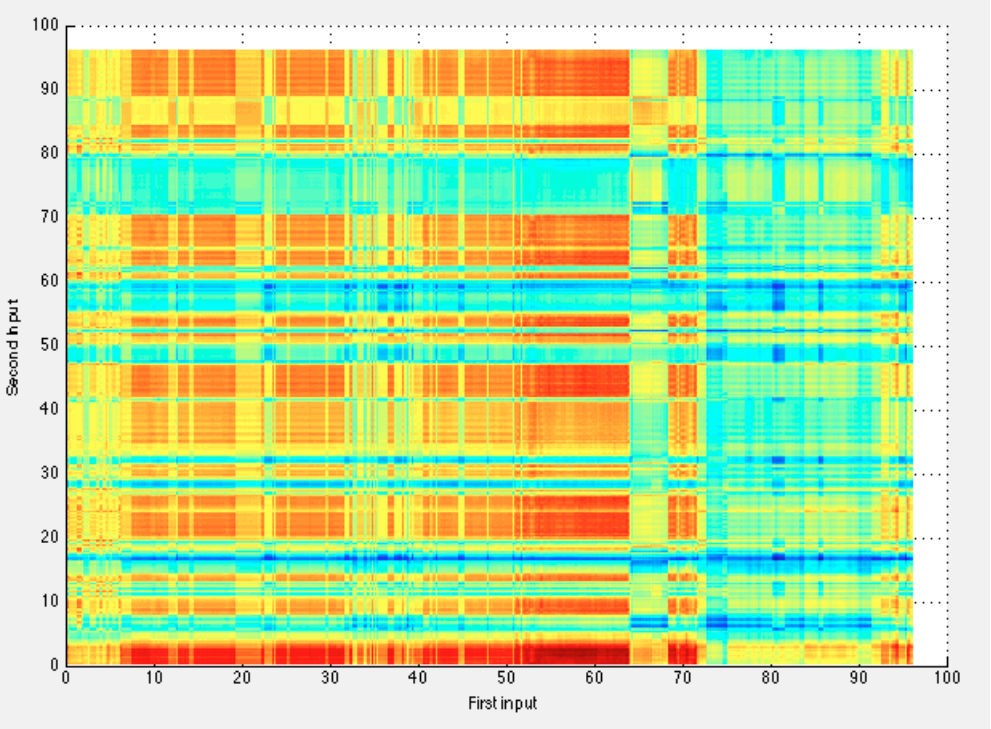
Figure 4.9. 3-dimensional recurrence plot



4.5.3.2 Cross recurrence plots

While Yang's (2010) gives an objective measure of posture similarity we also need to take into account the likely time course of mimicry. Posture similarity is unlikely to be instantaneous, copied postures may not appear simultaneously to the posture they intend to mimic but at some time delay to the original. For this reason a technique to uncover any recurrence of mimicked postures or bodily movements at different time intervals is required. Cross-recurrence quantification (CRQ) analysis is a nonlinear technique that exposes structure in complex systems over a continuous time series. The technique is known for sensitivity to space-time correlations and

Figure 4.10. 2-dimensional recurrence plot



determining recurrent states between two trajectories in reconstructed phase space (Shockley et al., 2003).

A cross-recurrence plot visualises recurrences of states in a dynamical system. For the purpose of plotting recurrence for posture similarity measures, similarity values for each frame are compared to every other frame in the file producing a three dimensional matrix of similarity values over the entire session. A MATLAB script to process the posture similarity analyses from the exported V files and produce the corresponding recurrence plots was developed by a programmer associated with the department of computer science at Queen Mary, University of London (Dr Chris Frauenberger). As can be seen in Figure 4.9, the recurrence plot affords a visual representation of this matrix, Figure 4.10 shows a two dimensional view of the plot. The colour represents the similarity value, where the resulting graph resembles a temperature map[6]. The least similar values are represented by blue, where the values become more similar as the hue shifts to green, yellow, orange, lastly red representing the highest similarity. The axes correspond to the two participants similarity value (labelled first and second input) in seconds, so the diagonal values in these plots represent the posture similarity value at zero

[6]The colours are relative to the range of values

offset (instantaneous), whereas at 0 seconds in the first input compared to the 50 seconds in the second input compares the first participants posture at 0 seconds to the second participants posture at 50 seconds. For example, in Figure 4.10 the band of red at the bottom shows the posture adopted by the participant entered as the second input in the first 3 seconds is similar to the posture adopted by the participant entered as the first input for over half of the interaction.

4.5.3.3 Measures of posture similarity

In addition to the 3-dimensional and 2-dimensional plots, the script also generated three overall measures of posture similarity for each pair of participants inputted, including:

1. Average overall recurrent posture similarity: As different people's posture never matches exactly a threshold level is required to determine what is similar enough to count as a recurring posture. This similarity threshold is set as greater than or equal to the mean similarity for that pair. This is a deliberately generous criterion designed to ensure even marginal cases of posture similarity are included. All posture similarity values over the threshold considered recurrent and averaged for this measure.

2. Determinism- the average proportion of repeated posture sequences: A percentage of recurrent diagonal sequences over a set time threshold. This was set to the average hand coded gesture length in this corpus: 2.5 seconds.

3. MAXLINE- the longest recurrent diagonal sequence: A recurrent sequence is any diagonally consecutive similarity values that are counted as recurring (greater than or equal to the similarity threshold).

4.6 Video-based analysis

Whilst Motion Capture technology is highly accurate and records useful specific data for analysis, the data cannot be fully representative of the unfolding interaction. To capture more of the detailed context several stages of hand-coding are used. Firstly, audio is captured as part of the video, which allows for transcription of the interlocutors speech providing the context and content of the nonverbal behaviour being studied. There is also pertinent visual cues which

are omitted from the motion capture data; fine bodily detail (such as facial expression, finger position and direction of gaze) as well as relevant features of the environmental setting with which the participants interact.

For an automated system to interpret behaviours, such as descriptive gesture, it be able to recognise the context, topic, as well as the movement characteristics that make up the gesture. The current state-of-the-art in automatic processing is unable to achieve this. A human coder is still the only practical approach to the annotation of complex behavioural events, such as descriptive gesture.

4.6.1 Aligning movement data to video

Video analysis was performed using ELAN[7]. Table 4.1 and Figure 4.11 describe the tiers created in ELAN for each set of movement data. Each tier is imported into the corresponding ELAN file and synced to the timeline along with the video. For nodding index both the cardholder and the non-cardholder each have a tier that contains a blank annotation for every instance of nodding on the timeline. For the hand movement data both the cardholder and the non-cardholder each have a tier that contains an annotation every 100ms that holds a numerical value of fastest hand movement measure at that point on the timeline (see Section 4.5.2 for more detail). The posture similarity data is represented in tiers that compare joint angles between interlocutors for various segments of the body specified in the table below. The similarity tiers for each segment also contains an annotation every 100ms; these annotations hold a numerical value of the corresponding measure at that point on the timeline (See Section 4.5.3).

4.6.2 Human coding and explorative observations

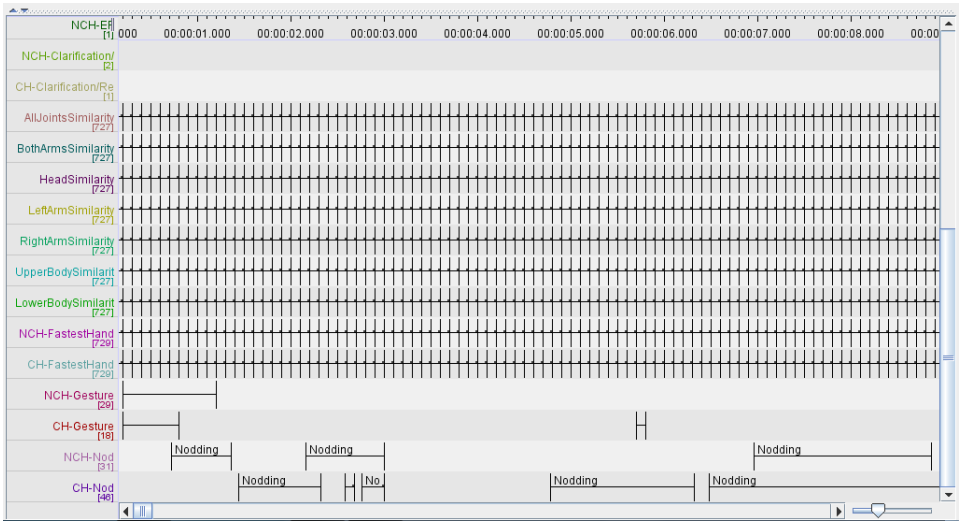
Video data is a useful for investigation as it visually (and audibly) preserves interactional events in time. This allows the behaviours and interactional conduct of the interlocutors to be analysed sequentially, with multimodal actions considered in relationship to one another. A qualitative video analysis of this corpus does not presume observations are indicative of usual patterns of

[7]An audiovisual annotation tool that allows multiple camera angles and audio to be played simultaneously, with the ability to align events as annotations on the timeline produced by the Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands: http://tla.mpi.nl/tools/tla-tools/elan/

Table 4.1. Motion capture data tier structure

Tier Name	Tier Description
AllJointsSimilarity	Value between 0-1 every 10ms
BothArmsSimilarity	Value between 0-1 every 10ms
HeadArmsSimilarity	Value between 0-1 every 10ms
LeftArmsSimilarity	Value between 0-1 every 10ms
RightArmsSimilarity	Value between 0-1 every 10ms
UpperBodySimilarity	Value between 0-1 every 10ms
LowerBodySimilarity	Value between 0-1 every 10ms
NCH-FastestHand	Value in mm/ms every 10ms for non-cardholder's fastest hand
CH-FastestHand	Value in mm/ms every 10ms for cardholder's fastest hand
NCH-Gesture	Annotation: Non-cardholder gesture
CH-Gesture	Annotation: Cardholder gesture
NCH-Nod	Annotation: Non-cardholder's nod
CH-Nod	Annotation: Cardholder's nod

Figure 4.11. Motion capture data tiers



behaviour, but is intended to undercover structures or mechanisms that are used in order to communicate.

Whilst human analysis of the video is contextually richer than automated processes, using human coders to interpret data accentuates the problem of *experimenter bias*- the idea that human coders tend towards interpreting events to affirm preconceived notions that they expect to see in the data while overlooking or even discounting counter-evidence (Rosenthal, 1963). This risk can be reduced by checking for inter-annotator reliability. A subset of the corpus is given to a second coder who annotates it independently. This is compared to check for consistency of annotation. The results for inter-annotator reliability is detailed in the introduction to the results section each analysis taken from video annotations forthcoming in Section 5.3, Section 7.3 and Section 7.3.2.

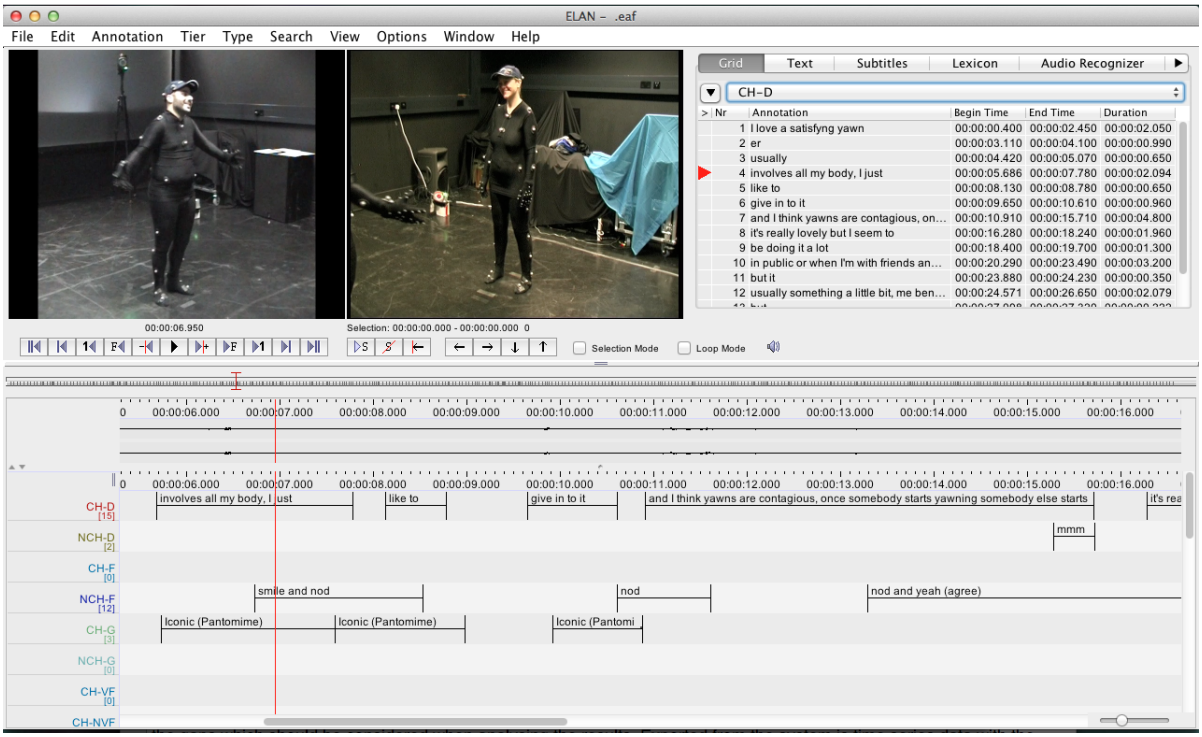
4.6.3 Preparing video for analysis

The first step of the coding process was to separate and organise the video into individual items. For corpus 1 the video and audio descriptions of each experience was separated into separate target items, for example a toothache is one target item and a yawn is another. For corpus 2 the dyadic interaction was separated out of the larger session. The video and audio was then imported into ELAN.

4.6.4 Transcription

Speech for each item was transcribed by the author, using ELANs waveform viewer as a guide for accurately aligning the utterance to the timeline. All verbal contributions from each participant were transcribed, including feedback speech such as 'mm hmm', 'uh huh' or interjections such as 'woah', 'ah ha' or 'er'. Laughter was not considered speech except for the when laughter was described verbally, such as 'Ho ho ho'. It was considered too verbose to align the transcript by word to the timeline so the transcript was aligned by utterance[8]. This preserved any pauses in-between utterances but for long contributions it is not possible to map individual words to a time within the utterance annotation (See Figure 4.12).

Figure 4.12. ELAN software environment for annotating video and speech



4.6.5 Gesture coding

As described in Section 3.2.7 there are different mechanisms people use to depict meaning. Normally categorised into five recognised gesture types: iconic, metaphoric, pantomime, deictic and abstract descriptive (McNeill, 1992; Rowbotham et al., 2011; Kendon, 2004). As this analysis looks at both cardholder and recipient strategies in formulating and communicating

[8]Utterance is defined as a string of words without any pauses between words

Table 4.2. Gesture tier structure

Tier Name	Tier Description	Controlled Vocabulary
CH-G	Cardholder Gesture	Iconic Metaphoric Pantomime Deictic Abstract Descriptive
NCH-G	Non-Cardholder Gesture	Iconic Metaphoric Pantomime Deictic Abstract Descriptive

their understanding of a description; a comparison of the patterns of gesture use between interlocutors is made. If mimicry is used in order to *simulate* the experience being described; topic gestures, the gestures representing the content of the experience being conveyed, are of particular interest.

4.6.5.1 Gesture timing

On the first pass, any occurrences of topic gestures were annotated without specifying their nature but ensuring the annotation fell on the correct point on the timeline. For this analysis there was only one annotation per gesture type.

The gesture starts as soon as the participant starts to move into place, for example moving a hand from by their sides to their gesture space. This preparation stage counts as part of the duration of the gesture. The same goes for retracting the gesture, the hand movement from the gesture space back to their side again is also counted as a part of the gesture.

Including the preparation and retraction stages in the gesture annotation can present potential inconsistencies in gesture length for each annotated gesture. Sometimes, a participant will move

from one gesture type to another without dropping their hands. In this case the preparation stage of second gesture can overlap with the retraction stage of first gesture. In this case what would have been the retraction movement of the first gesture is annotated the second gesture preparation stage, there is no overlap. This usually happens very quickly but does produce the effect of cutting the first gesture short, creating an inconsistency in length. Counting overlaps like these would inflate the total gesture time in the analysis.

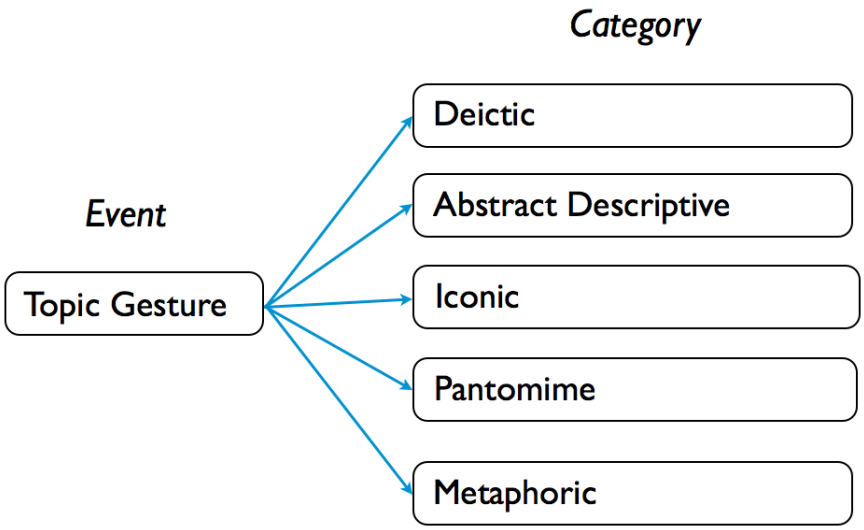
Holds[9] are counted towards to duration of the gesture. For example, if a participant points to their tooth and holds their point until they have finished their utterance, then the hold counts towards the duration of the gesture, here the annotation continues until the participant either retracts the pointing gesture or moves into another gesture. Holds outside of turns do not count towards the duration of the gesture. For example, if a participant is speaking and pointing, then the recipient interjects and the participant stops speaking in order to listen, this does not count toward the duration of the gesture. Holds outside of turns are very rare in these corpora and usually the speaking participant lowers their hands a little while holding then raises when they return to their speech. These holds are not of interest for this process of gesture annotation as they are not intended convey information to the recipient.

4.6.5.2 Determining gesture type

On the second pass, topic gestures were separated into five types: Iconic, Deictic, Pantomime, Metaphoric and Abstract Descriptive. The annotator followed the definitions indicated in the gesture taxonomy (see Table 4.2 and Figure 4.13). It is important to mention that all topic gestures were coded irrespective of whether the cardholder or non-cardholder performed the gesture, so both participants were coded for their gesture and the same definitions were used for each. To reduce experimenter bias, on each pass participants were coded one at a time; with the viewer only displaying the participant that was currently being annotated. This was achieved by either only displaying one video stream for corpus 1, or covering one participant by placing a piece of card on the screen for corpus 2.

[9] A gesture phrase consists of one or more movement phases. The stroke phase is the section of movement that expresses the meaning of the gesture. The stroke phase is preceded and followed by a preparation and retraction phase. A *hold* is the temporary pause in movement without leaving the gesture phrase, holds can occur before or after the stroke phase (pre-stroke hold / post-stroke hold) (McNeill, 1992; Kita et al., 1998).

Figure 4.13. Gesture taxonomy



Very occasionally a gesture will cover more than one of the gesture types. For example a participant is pointing at their tooth, then with the same hand starts to produce a motion describing how it feels whilst still pointing. Counting one gesture as two different types inflates the total number or duration of gestures in the interaction. In attempt to avoid this the overlapping gesture types were split where this was possible. For example if a gesture starts as a pointing gesture, then describes a motion while still pointing; a Deictic gesture would be annotated followed by an Iconic gesture when the motion gesture starts. This has the limitation of disregarding the dual nature where the Iconic gesture overlaps the Deictic gesture. If impossible, the gesture was annotated for the most dominant form, exercising judgement as to which gesture is intended to convey more information than the other at a particular time. There are problems with this way of annotating so to check reliability a sample of randomly selected items was independently coded by a second coder and reliability measures (Cohens Kappa) calculated.

From this process every instance of topic gesture, its type, whether the cardholder or recipient performed it, and for what target item it belonged was recorded and prepared for statistical analysis.

4.6.5.3 A taxonomy of descriptive gestures

For current purposes a systematic taxonomy is required that allows a discrimination between gestures of functional difference. A review of literature categorising different types of gestures within an interaction follows, seeking to build a taxonomy in which a suitable context for interactions centred around descriptions of felt experience can be placed. Gesture, for current purposes, includes anything nonverbal[10] that is produced as part of the conversation that includes hand gestures or body movements and facial displays that descriptively enact the content of the talk.

Of interest are the narrative elements of an interaction, these refer to gestures that depict or are about the content of the speech[11] (Bavelas et al., 1992). These are spontaneous gestures that occur mostly synchronously with speech, also referred to as as physiographic (Efron, 1941) or lexical movements (Krauss and Chen, 1996).

In order to differentiate between narrative gesture types, this work uses a similar subdivision of gesture categories to McNeill (1992): *Iconic*, *Metaphoric*, *Deictic*[12] and supplements it with the *Pantomime* category described by de Ruiter (2000) as well as the *Abstract descriptive* category described by Rowbotham et al. (2011).

Pantomime gestures are defined as imitations of functional motor activities, body enactments that are generated from 'procedural motoric knowledge' (de Ruiter, 2000, p. 294). In interactions that involve story-telling, a pantomime gesture can give a particular type of insight

[10] nonverbal is used as a gloss to mean non-speech communicative actions such as conversational gestures and body movements.

[11] For this work, gestures that do not contribute to the content of the message are excluded. These include adaptors- hand movements such as head scratching or directed towards objects- such as drinking out of a glass, and gestures that form para or meta narrative elements of interaction. Para-narrative gestures function to manage the interaction itself. Bavelas et al. (1992) terms these *interactive gestures* as they help to maintain conversation as a social system- these can be used to refer to an recipient, cite them, seek response or confirmation of understanding, emphasise words, turn taking and so on. Meta-narrative elements are gestures that depict aspects of the delivery of talk, emphasising the certain words or iconically describing the flow of the speech (such as beats or batons).

[12] Excluded is *Emblems*, otherwise known as conventionalised gesture (Gullberg, 2009) or symbolic gesture (Ekman, 1969; Krauss and Chen, 1996). These have a known meaning across a culture and are independent of speech, although can accompany it. These were excluded as they lack the spontaneous descriptive characteristic that focuses on the features of the experience.

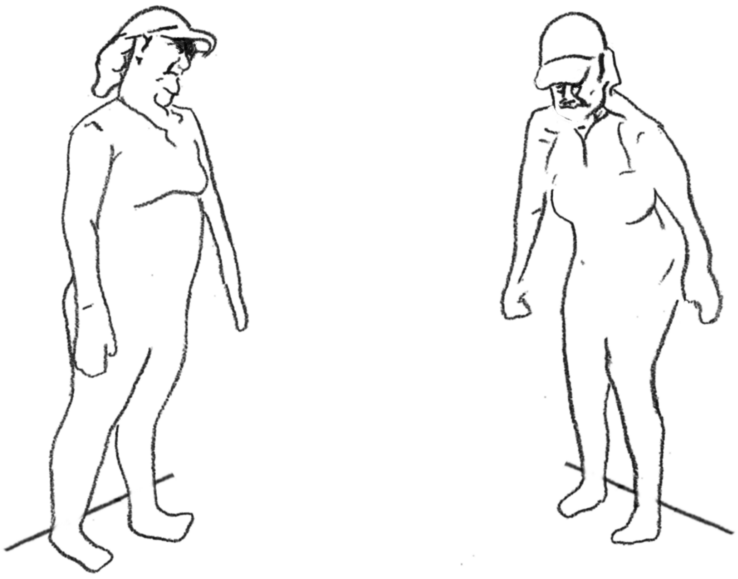
into a situation. For example, when describing an experience of stubbing a toe, a storyteller may hop around on one foot depicting their outward behaviour caused in reaction to stubbing their toe. This re-enactment of their behaviour is intrinsic to the original action and so would be classed as an iconic gesture by McNeill (1992), however because of the emotive context pantomimes provide, this thesis separates out pantomime gestures after de Ruiter (2000) for further analysis.

In the following example of a pantomime gesture, the participant on the right of Figure 4.14 labelled CH is describing her reaction to the pain of a backache. She accompanies her speech in lines 1 and 2 'I ended up going- leaning forward' by hunching her back, shoulders tense around her ears, her face in the shape of a grimace. As if she were performing the tense movements caused by the pain. She then says in line 3 'and I walk around like I'm ni:nety years old!' while bending but stiffening her legs as if she didn't have the strength or flexibility to straighten them, then producing a walking movement, tensely moving one foot out and alternating if if she were shuffling forward on the spot. This exhibits an enactment of the reaction to an experience, exaggerating the movement for maximal effect.

In the next Figure 4.15, she elaborates the pantomime by saying in line 5 'and it's just like- "Oh my god"- I'm not even the:re yet and I'm already like a::rgh!' while returning to the crouched over posture. In Figure 4.16 she says 'all crunch all arms down and crunch and it's just an awful pain' in line 6, this time adding arms by putting her arms out to her sides, bent but rigid, and again performs almost robotic shuffle forward to indicate her range of moment had been hindered by the pain. Her speech forms the commentary of the pantomime gesture that enacts her reaction at the time of experiencing the pain.

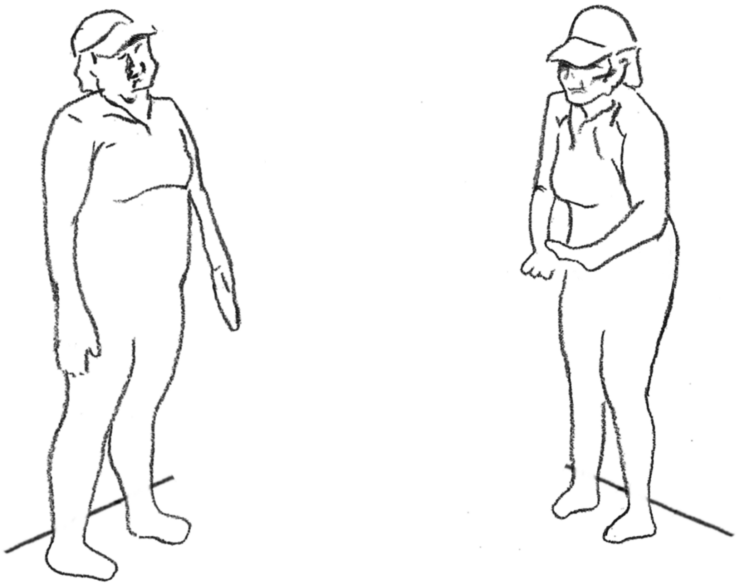
Pantomime gesture do not always have to be enacted from one's own point of view but can be an enactment in another perspective. Pantomime gestures nonverbally demonstrate an expressive reaction to an experience, event or situation, regardless if the performer had the experience or not.

Figure 4.14. Leaning forward sketch



1. CH: I ended up going
   (0.4)
2. CH: leaning forward
   (0.8)
3. CH: and I walk around like I'm ni:nety years old!=
4. NCH: =Yeah
   (0.3)

Figure 4.15. ... I'm not there yet sketch



5. CH: and it's just like

(0.4)

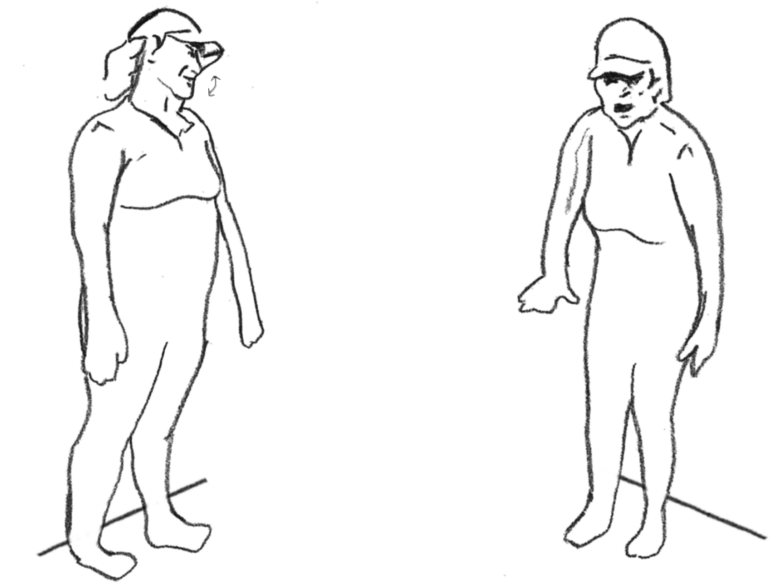
6. CH: Oh my god I'm not even the:re yet and I'm already like

(0.3)

7. CH: a::rgh!

(0.4)

Figure 4.16. Crunch sketch



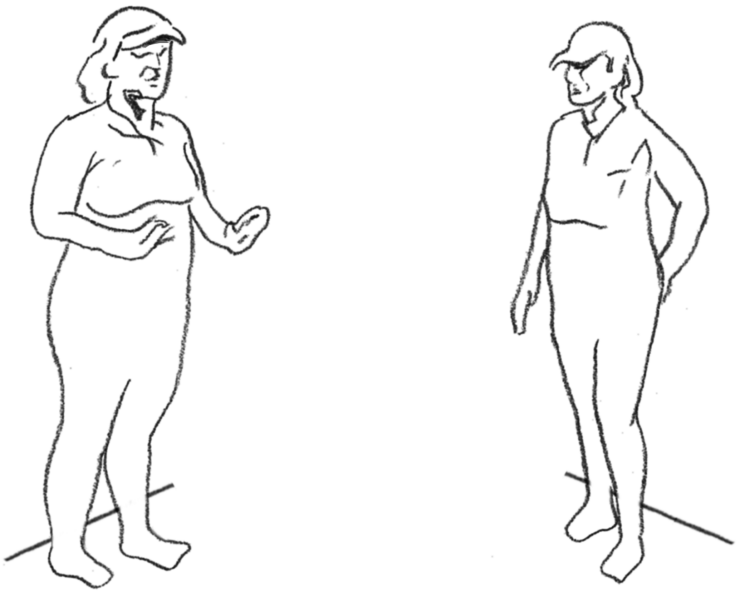
8. CH: all crunch all arms down and crunch and it's just an awful pain=

9. NCH: =[Yeah]

[[nods]]

(0.6)

Figure 4.17. Don't want that sketch



10. CH: an awful pain=

11. NCH: =No I'm glad I'm done with that, I don't want it to come back=

12. CH: =That's good, yes.

Deictic gestures

This is a gesture that points to the object of talk. A simple example would be when an interlocutor points to something concrete, like an object or person in the room, because they are talking about it and wish to locate it for the recipient.

In the following example sketched below in Figure 4.18 the participant labelled CH says in line 1 and 2 'Toothache. Actually I've never really- had one. Oh yeah maybe.' In the sequence he points towards his mouth with his right index finger twice, once when he says the word 'Toothache', pointing at his mouth from the right side. The second point, although performed again with the right index finger, points more towards the centre of his mouth, using his gesture to point to the specific site of the experience he is describing.

Deictic gestures can be used to locate something that cannot be physically pointed to but referred within the speakers' gesture space. In these cases a speaker has constructed the spatial location of the object they are referring to in their gesture space, referring to an house on their left and a church on their right to describe a concrete location of the object of their speech.

Consider the following example, the participant labelled CH is describing the layout of her office at work. In Figure 4.19 she describes where she sits by saying 'Like at work 'cos the lighting is really bad and I sit in like this sort of dark corner off the side and the light over me is broken' in line 1. At the same time she utters 'dark corner off to the side' she twists her body and places her hands to the left of her left hip. Indicating in her gesture space that she sits right out of the centre of her office but in a 'dark corner off to the side', here she uses space in front of her that she would usually gesture in to refer to the central point of the office and compares the uses it to offset where her desk sits in relation to that.

In Figure 4.20 she says 'And umm- and so I've got this really horrible desk lamp which is like- Th:is tall' in lines 2, 3 and 4 while pointing up. She indicates the height of the desk lamp without explicitly pointing to a location, for this reason this gesture would not be considered as a deictic gesture, but as it describes concrete aspects of an object it would be regarded as a *iconic gesture* which is described in more detail below.

She continues in Figure 4.21 by saying 'And the light either shines directly in my face or directly at my colleague just across from me and it's like really annoying and horrible and er so I get

Figure 4.18. Toothache? Sketch courtesy of Daniel Adderley

1. CH: Toothache. Actually I've never really
   (.)
2. CH: had one. Oh yeah maybe.

these headaches everyday because it's either too dark if I have the desk lamp off or too bright with the desk lamp' in lines 5, 6 and 7. At the moment she says 'my colleague over there' she points a metre or so in front of her, this time she is using her own body as her perspective at the desk and points in front to describe where her colleague sits in relation to her.

Figure 4.19. I sit here sketch

1. CH: Like at work 'cos the lighting is really bad and I sit in like this sort
   of dark corner off the side and the light over me is broken
   (0.4)

As we have seen pointing to the location of what is being referred to can be anchored from different perspectives, the speaker creatively using the resources they have to describe points locational information.

Deictic gestures are can also refer to something in the abstract (McNeill et al., 1998). Often a speaker will construct abstract ideas spatially in their gesture space. For example, when a speaker says 'on the one hand' and points to one side of their gesture space.

Figure 4.20. This tall sketch

2. CH: And umm

(0.2)

3. CH: And so I've got this really horrible desk lamp which is like

(0.7)

4. CH: [Th:is] tall

[(reaches up)]

(.)

Figure 4.21. Colleague over there sketch

5. CH: And the light either shines directly in my face or directly
   [at my colleague just across from me] and it's like
   [[points]]
   (.)
6. CH: really annoying and horrible and er so I get
   (0.3)
7. CH: these headaches everyday because it's either too dark if
   I have the desk lamp off or too bright with the desk lamp

Iconic gestures include a depiction that is intrinsic to the content that is being conveyed, for example a gesture describing a ball will depict a characteristic intrinsic to a ball, such as making a fist to represent roundness by making the hand round. Iconic gestures are always a concrete entity or action rather than abstract or analogous. Perhaps more relevant is iconic gestures that act out being in pain or the cause of the pain. For example, when describing an experience of stubbing a toe, the speaker may hop around on one foot depicting the outward behaviour caused by stubbing their toe. This mimicry of their behaviour is intrinsic to the original action and so therefore are classed as a iconic gesture. They encode a speakers viewpoint on a communicated depiction. Ekman (1969) include deictics, spatial, kinetographs, pictographs and rhythmic elements in this category. Beattie and Shovelton (2002) suggests that the content of these gestures to describe action, shape, size, direction, speed, and relative-positions.

In the example below shown in Figure 4.22, the participant labelled CH says 'they got- they got the biggest pair of pliers I've ever seen' in line 1 while he puts his hands in front of time, approximately 40cm apart, palms facing in to describe the size of the pliers he refers to in his speech. The distance between his hands refers to the length of the pliers. He does this horizontally, so it could be said to be ambiguous as to whether he is referring to the length of the scissors, the width of the pliers or whether they are open or closed. Despite the ambiguity, it is clear that the distance between his hands refers to a length because of the reference to size in his utterance when he says 'the biggest pair of pliers' negating the need for his to verbally describe the exact size in words because simultaneously to his speech his hands provide this information. The size of an object is a concrete characteristic, because of this the gesture would be categorised as an *iconic* gesture.

Figure 4.22. This big sketch

1. CH: they got- they got the biggest pair of pliers I've ever seen

Metaphoric gestures: are usually pictorial but unlike iconic gestures, metaphorics depict abstract ideas rather than a concrete object or event. The topic of the metaphor being the abstraction, the vehicle or gestural image being the offered virtual object spatially localised, and the common ground of meaning is where the topic and vehicle are linked in properties, such as physical containers. For example, in a description of a backache an interlocutor describes the most extreme sensation of that particular pain by saying 'and that was the crest of it', while reaching high and pointing to the top point in the gesture space, spatially highest gesture metaphorically signifying the most intense pain.

Abstract descriptive gestures Rowbotham et al. (2011) make an additional subtype of topic gesture they term abstract descriptive, these gestures describe inner experiences and are categorised by featuring 'imaginist and semantically related to speech but containing information which could not be visually accessed (subjective experience of pain), therefore not iconic according to McNeill. This was after Hyden and Peolsson (2002), who observed this category of gesture as constructing pain as an 'object' in the gesture space, gestures that were used for describing the pain or some aspect of the pain, not including pointing to the location of the pain or miming the bodily effects of the pain. In other words, these gestures describe the felt sensation without acting it out, and express more than just the location of the sensation. They specifically refer to the inner sensation of the experience, put plainly, what the sensation

feels like. For example, when describing the sensation of stubbing ones toe we might depict the pain with a gesture that uses our hands to represent the rhythmic quality of a throbbing pain by mimicking it in the rhythm of our hand movement, or perhaps the intensity intrinsic to the pain would be depicted by the tenseness of our fingers. This type of gesture does not represent an analogy of the felt experience so cannot be a metaphoric gesture but on the other hand cannot be symbolised concretely, as the felt experience is only accessible to the experiencer. This indicates that they lie somewhere on the borders of iconic gesture and metaphoric gesture.

The example below show the stages of two abstract descriptive gestures, relating to the sensation when at of the onset of a fit of laughter. The participant breaks the experience down into two parts, the first part is shown in Figure 4.23 and the second in Figure 4.24. The first part the participant, again labelled CH says in lines 2, 4 and 5 'like a sort of a- fountain- like an eruption!', he uses two different descriptive words but repeats the same gesture three times when he says 'like a sort of', the word 'fountain' and the following utterance 'like an eruption!'. The participant labelled NCH is his interactional partner but is omitted from the sketch, however their utterances are still captured in the transcript for the sake of faithfulness to the interaction as it unfolded. Here the unseen participant interjects over the end of the utterance in line 2 with an affirmation of previous talk in line 3 when they interject with 'a:h you were fifteen'. The participant in the sketch draws a curve from his waist hands starting point downwards, then moving inwards and rotating his palm to face one another up to the bottom of his chin mirroring the action with both hands, trajectory moving upwards and inwards, he pauses slightly at the top then resets his hands back down outwards to the starting point again to begin the gesture again.

The second part is shown in the second sketch, which is only performed once during line 7 'and then a release', his hands are now at the bottom of his chin pointing inwards, he than rotates them palms facing outwards and pushes forward quickly then when he reaches the breadth of his movement stops momentarily and drops his hands downwards, slightly out, his hands seemingly falling with gravity. He swings his arms back and forth slightly at the moment his interactional partners utterance 'a:h yeah I guess so', perhaps to indicate his arms are now relaxed and agile.

Although the word 'fountain' could be attributed to a concrete object the way it is used in this

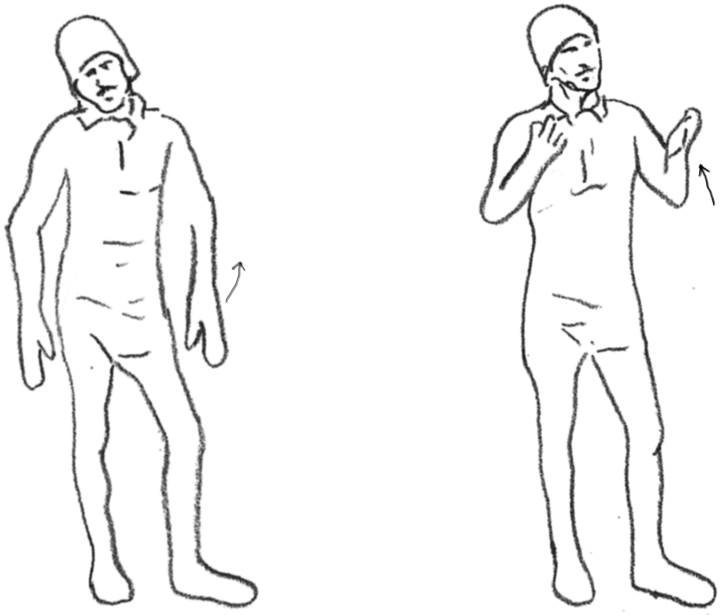
example is dynamic and seems to refer to the movement of water being sucked into a water fountain, used metaphorically to mapping the dynamic way the water is sucked in and the pressure it holds at the point before it's released again to the feeling of increasing intensity and the pressure that will eventually overflow. He uses this dynamic quality of the metaphor with the trajectory and velocity of his hands to describe is visually, the location of his hands could be clues to whereabouts the sensations moves. This use of metaphor to describe a feeling state- something that is not a concrete thing and cannot be directly shared between him and his interactional partner, for this reason it is referred to as an abstract descriptive gesture. The repetition of the gesture to his utterance 'like an eruption!', uses instead of the dynamic quality of a moving object (or fluid), the dynamic quality of an event or instance. Again he metaphorically maps the movement that the words 'eruption' describes, (defines in the Oxford English dictionary as 'an act or instance of something exploding' or 'a sudden outbreak of something, typically something unwelcome or noisy'), the movement of something suddenly moving outward as the result of built up pressure. Here he creatively chooses another metaphor but uses the same gesture to visually describe the dynamic build up of sensation. This again would be categorised as an abstract descriptive gesture as the metaphor describes something abstract- the sensation of 'bursting into laughter'.

The last part of the sequence sees the participant push his hands out from just in front of him to about 50 cms in front, this occurs over the utterance in line 7 'and then a release'. Here he uses himself as an anchor for the origin of the thing that is released and pushes out in the direction he deems the thing released goes. Here as he is talking about laughing, the movement could correspond to a the act of laughing producing a sound that originates in the mouth, as laughs do. He could also be marking the ceasing of a building sensation as moving out of the body with force. Again, the gesture is typically metaphorical in form, buts refers to an abstract thing, a sensation.

4.6.6 Feedback behaviour coding

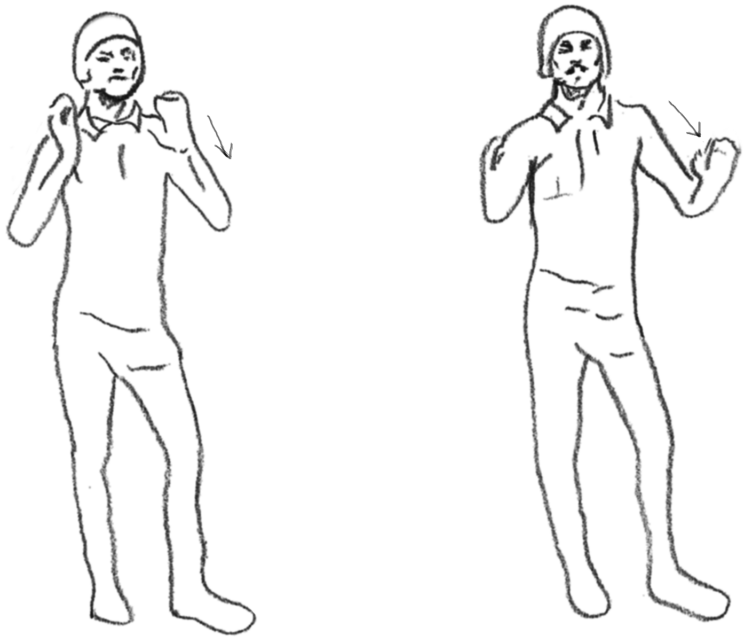
As discussed in Section 3.2.8, feedback behaviours are an important interactional resource to a constitution of intersubjectivity. Gaining a better understanding of how and where feedback is used to demonstrate an understanding of what each other experiences and to what effect; is the purpose for collecting data on feedback behaviours.

Figure 4.23. Fountain laugh sketch



1. CH: for a fifteen year old and yeah it was just
   (1.0)
2. CH: like [a sort of a]
3. NCH: [a:h you were fifteen]
4. CH: [fountain]
   [(raises hands)]
   (0.4)
5. CH: [like an eruption!]
   [(repeats raising hands)]
   (.)
6. NCH: Yeah
   (1.3)

Figure 4.24. ...and release sketch



7. CH: and then a [release]=

   CH: [(pushes hands forward)]

8. NCH: =a:h yeah I guess so

Figure 4.25. Jabby toothache: Sketch courtesy of Daniel Adderley

1. CH: It was really-

   (0.4)

2. CH: it was painful

   (0.5)

3. CH: er very se::aring pain

   (0.2)

4. CH: very very

   (0.4)

5. CH: jabby!

   (0.4)

6. CH: you know? Like

   (0.5)

7. CH: almost as if someone stabbed you. [Obv]iously I hadn't been sta[bbed]

   NCH: [Mmm] [Yeah]

In order to analyse the character of feedback behaviours this analysis separates feedback instances by modality. Each participant has two feedback tiers; verbal or nonverbal. Each instance of feedback was then coded according to it's function within the interaction. There are three recognised functions of recipient feedback:

- *Contact and Perception (CP)*, indicating recipient contact and perception of message (Allwood et al., 1993; Cerrato, 2002).
- *Comprehension (C)*, indicating recipient comprehension or understanding of message (Allwood et al., 1993; Cerrato, 2002).
- *Attitudinal or Emotional (A/E)*, indicating an attitudinal or emotional response as simple as agreeing with the speaker (attitudinal), or showing shock to the message (emotional), like motor mimicry (Cerrato, 2002; Schröder, 2003; Heylen, 2006; Bavelas et al., 1987).

Table 4.3. Feedback tier structure

Tier Name	Tier Description	Controlled Vocabulary
CH-F	Cardholder Feedback	None (Noted Observations)
NCH-F	Non-Cardholder Feedback	None (Noted Observations)
CH-VF	Cardholder Verbal Feedback	CP (Contact/Perception) C (Comprehension) A/E (Emotional/Attitudinal)
NCH-VF	Non-Cardholder Verbal Feedback	CP (Contact/Perception) C (Comprehension) A/E (Emotional/Attitudinal)
CH-NVF	Cardholder nonverbal Feedback	CP (Contact/Perception) C (Comprehension) A/E (Emotional/Attitudinal)
NCH-NVF	Non-Cardholder nonverbal Feedback	CP (Contact/Perception) C (Comprehension) A/E (Emotional/Attitudinal)

These functions display different levels of engagement and understanding from the recipient.

They are also cumulative, so by producing feedback demonstrating comprehension (by uttering 'I understand' or 'I get it' for instance) would naturally also show contact and perception, as interlocutors would have had to show they are paying attention in order to that they understand. If an recipient laughs at an amusing description of a something, they are demonstrating an emotional response, but this feedback encompasses the function of showing comprehension; *and* contact and perception. So the highest function level overlaps the lower ones. In this analysis these levels are separated without overlaps, the feedback is annotated as the highest level function it demonstrates. For example if an interlocutor is laughing at the same time as nodding; this instance would be annotated as emotional or attitudinal (A/E) feedback.

There are some issues with annotating feedback in this manner. As each instance of feedback has one annotation it is sometimes difficult to see where one instance stops and another begins, which is particularly applicable when detecting pauses in-between nodding. Also an interlocutor could nod for a long period of time which would produce a lot of feedback but will only be counted as one instance of feedback; it is important to consider this effect if quantifying feedback by counting occurrences.

Sometimes feedback is given by making a remark about the talk they are party to. Take this example- a participant is giving a description of their experience of a recent backache when their recipient interjects with a series of comments about the backache, the recipient then goes on to talk about their own backache. It is sometimes difficult to judge when recipients go beyond remarking on the talk and developing the topic. This extension is not regarded as recipient feedback for the purposes of this analysis.

4.7 Identifying repair

4.7.1 Clarification questions

A clarification question is asked when a recipient fails to fully comprehend something in their interactional partners previous utterance or contribution. The form of a clarification question (otherwise known as a clarification request (Purver et al., 2003) or second position repair initiations) can differ vastly depending on what the recipient believes they don't understand or misheard from the previous talk and the manner in which they attempt to resolve the misun-

Figure 4.26. Feedback function categories

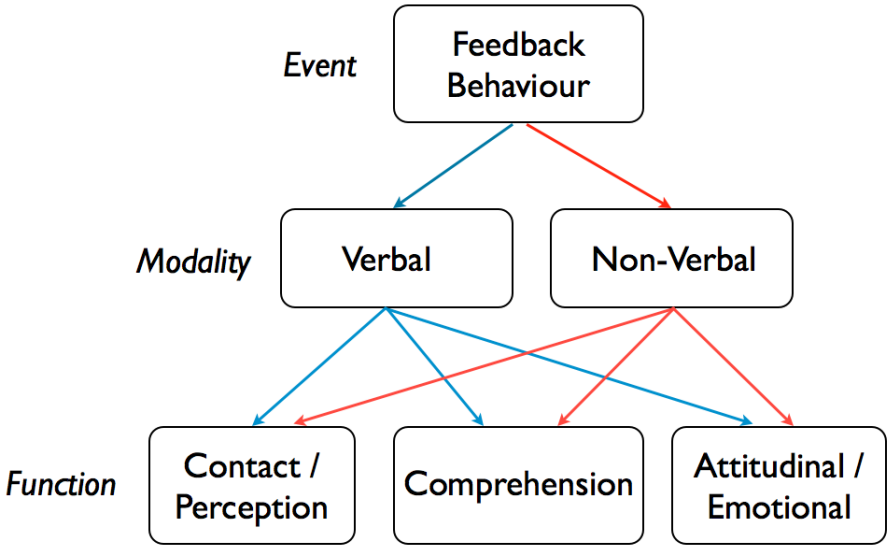


Table 4.4. Clarification question/repair tier structure

Tier name	Tier description	Controlled vocabulary
CH-Clarification/repair	Cardholder Clarification/Repair	CQ (Clarification question) R (Response)
NCH-Clarification/repair	Non-cardholder Clarification/Repair	CQ (Clarification question) R (Response)

derstanding. For this reason it is considered more practical to hand annotate the corpora for any occurrences of clarification questions and their responses rather than attempt to build an automatic system.

The annotation of Clarification/Repair was conducted in ELAN where each participant has a Clarification/Repair tier and each instance of repair of this type are annotated onto the timeline from analysis of the video. As this type of repair nearly always occurs as an adjacency pair (question and answer), each instance is separated into two parts. Annotations are labelled either CQ (Clarification Question) or R (Response) depending on whether the participants contribution forms the question part or the answer part, although occasionally clarification

question are left unanswered (See Table 4.4).

Robust annotation of the response section is sometimes difficult as the timing and form of the response can vary from the typical question immediately followed by an answer format. Firstly, the response to a clarification question is not always verbal, sometimes they can be in the form of a nod or a point toward the recipient to confirm they are correct. Nonverbal response contributions are labelled as normal responses (R). Secondly, occasionally the question is answered before the initiator finishes formulating the question. This occurs when the remainder of the question is anticipated from the opening words of the question. This results in an overlap of question and response. The full contribution of each clarification question and response is annotated regardless of overlap. Lastly, it is sometimes difficult to judge when the responder has finished answering the clarification question and continues the original talk. For the purposes of this analysis the duration of the annotation encompasses the response contribution without elaborating into something irrelevant to the clarification question.

4.8 Statistical significance

Following standard practise, this thesis uses a criterion level of p = <0.05. Only P values less than 0.05 are considered significant.

Part III

Mimicry in interactions about bodily experience

Chapter 5

Gestural mimicry

5.1 Introduction

As mentioned in Section 2.3.1 on emotional contagion, studies analysing facial mimicry find that that people nonconsciously mimic certain observed facial expression at a pre-attentive level. It is suggested that mimicry of facial expressions allows for an inner simulation of the affective state associated with an expression due to overlapping motor, sensory and affective neural representations. However, according to Bourgeois and Hess (2008); Iacobini et al. (2009, 2010), not all expressions were found to be mimicked, for example, expressions of anger are responded to with neutral expressions, this suggests that only expressions where mimicry is an appropriate affiliative response are responded to with a matching expression- reflecting angry expressions does not signal affiliative intent but a threatening display. Studies on facial mimicry are based on responses to photographic or video-based stimulus and not during interaction with another interlocutor, also mimicry responses were considered when a matching expression is observed whilst perceiving a stimulus expression or instantly after.

The majority of studies on behavioural mimicry as a chameleon effect detailed in Section 3.2.1 do measure mimicry within interaction but their definition of mimicry operates a different temporal restriction. Mimicry is considered as matching behaviour throughout an entire interaction and is based on the frequency of mimicry behaviours rather than whether an expression is simulataneously matched. These studies take their analysis from interactions with confederates- who either purposely mimic participants behaviour; or produce behaviours, such as foot-tapping

or face-touching, that have subsequently been observed in the participants behaviour over the course of the interaction (see Section 3.2.1). In both variants there are certain aspects of the confederates behaviour not spontaneously produced as a component of the interaction (but for the purposes of detecting mimicry). It is unknown what the effect of this would be on the structure of the resulting interaction or on behavioural mimicry.

A study analysing behavioural mimicry that did not use confederates is Griffin et al. (2015) work on gestural mimicry during laughter in triadic spontaneous social interactions. Both laughter and mimicry are claimed to be rapport building social signals and so Griffin et al. (2015) suggests as they functionally overlap one another, there may exist a causal relationship between the two behaviours. This study did place a temporal restriction on what was considered mimicry- defining an exact matching within 3 seconds of the original behaviour as mimicry. They found that although participants produced more gestures when laughing, mimicry did not occur more frequently during laughter episodes. However, mimicry did increase in the second half of the session, suggesting that more mimicry occurred as rapport was built throughout the interaction.

Analysing mimicry of foot-shaking and/or face-touching also has the limitation that these mannerisms do not contribute to communicating the primary message of the talk- although perhaps an insight into the (albeit spurious) stance of the confederate producing the face-touching or foot-shaking or emotions related to what is being communicated. Although there are studies that look at gestural mimicry for content-specific gestures (see Section 3.2.7.1) these do not measure whether there is an overall matching of descriptive gestures between interlocutors due to a perception-behaviour link that leads to automatic mimicry.

This chapter analyses interactions between two participants with a goal to share and understand one another's descriptions of recalled bodily experiences, a situation expected to promote empathetic interactions. It is anticipated these interactions focussing around retellings will feature descriptive gestures contributing to conveying, or responding to, an account of the bodily experience. The analysis a similar approach to measuring mimicry to the studies that investigate the perception-behaviour link by counting the frequency of behaviours over the duration of an interaction, no time restriction is placed on what is counted as mimicry. In this study, constrastingly, it is not specific behaviours such as foot-shaking or face-touching that are counted, but descriptive gestures are separated into categories and measured to obtain an overall sense if

there is the potential for mimicry to occur within these distinct categories (see Section 4.6.5.3 for categories).

It is expected that, as descriptive gesture is related to the message being conveyed, if mimicry is a vehicle to understanding of another's affective experience from embodying a person's expressions in ourselves, then these descriptive gestures would be crucial component of the described experience to embody. If empathy have simulative underpinnings, such as those described in the perception-behaviour link, then bevavioural mimicry would extend to descriptive gesture.

Drawing from Corpus 1 (bodily experience dyads) as laid out in Section 4.2 comprising of a collection of dyadic conversations about six different bodily experiences. This section presents an analysis of gestures produced by the pairs of interlocutors for each experience type. For analysis of experience type, target items were separated intuitively into either a 'Positive' valence category (Satisfying yawn, Laughing out loud, Back massage) or a 'Negative' valence category (Backache, Stomach ache, Toothache). Two task roles are distinguished: 'Cardholder' for the participant currently relating their experience on their card and 'Recipient' for the participant who is audience for the description. The analysis considers descriptive gestures expected to directly contribute to conveying an account of the participants recalled bodily experience: Iconic, Metaphoric, Abstract descriptive, Deictic and Pantomime gestures.

The patterns of gesture use found in Corpus 1 are then compared with Corpus 2 (control dyads) as laid out in Section 4.3, comprising of a collections of dyadic conversations about current popular topics. A comparison across both corpora helps to clarify what different gesture types contribute to descriptions of felt experience in comparison with other conversational topics.

5.2 Predictions

If participants are imitating one another's gestures, the imitated gesture (the stimulus) and it's imitation (or imitations) would have to match in categorical type; for example, if you copy a pantomime gesture you would have to produce a matching pantomime gesture- so both gestures are pantomime gestures. A simple measure to detect a similarity between interlocutors is to compare the pattern of gesture type distribution for each conversational partner. It would be

also assumed that matching gestures would be of a similar length. So the duration of gesture types are also considered. Following from the hypothesis that a perception-behaviour link or inner simulation forms the basis of behavioural mimicry- if interlocutors are automatically and systematically mimicking one another's gestures; two base predictions for gestural mimicry will follow:

1. The distribution of gesture type occurrences and length would match between cardholder and recipient.

2. If there is a difference in the distribution of gesture type occurrences and length for descriptions of positive and negative valence, both cardholder and recipient gesture pattern would reflect this.

5.3 Results

Inter-rater reliability for the five descriptive gesture categories was very good with Kappa = 0.702, p <0.001, n = 68, Standard Error = 0.06.

5.3.1 Overall gesture type distribution between cardholder and recipient

As Figure 5.1 shows the overall distribution of gesture types produced by people when they are in the two roles is quite different. Table 5.1 shows Recipients produce far fewer descriptive gestures than Cardholders (15.7% vs. 84.3% of overall descriptive gesture).

Figure 5.1. Incidence of descriptive gesture

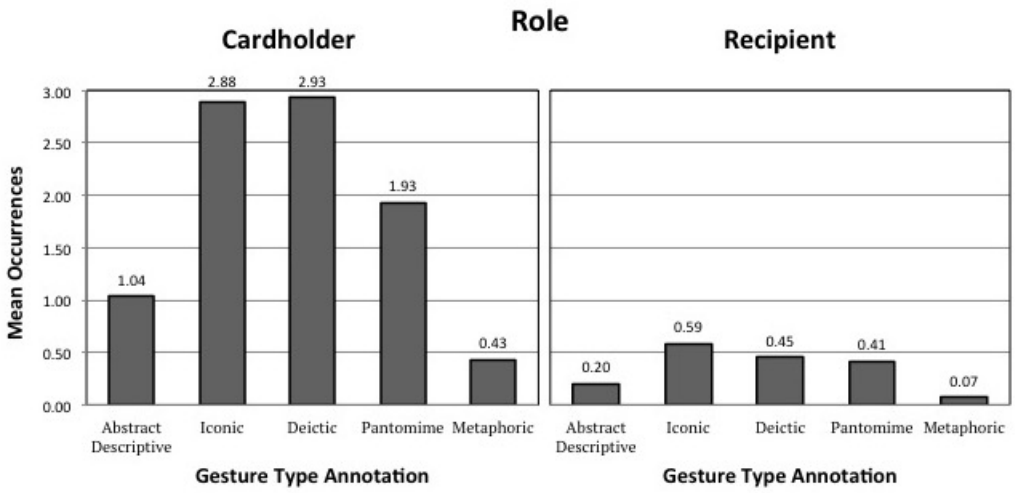


Table 5.1. Total gesture occurrence for cardholder and recipient

	Cardholder	Recipient
Total gesture occurrence	1022	191

Chi^2 comparisons show within overall pattern of use for Iconic, Deictic, Pantomime and Metaphoric gesture is not reliably different for Cardholder and Recipient ($Chi^2_{(3)}$= 1.94, p = 0.41). However, a focussed (post-hoc) comparison shows the use of Abstract descriptive gesture differs comprising around 10% of Cardholder's gestures but only 1% of Recipient's gesture ($Chi^2_{(1)}$= 95.6, p <0.001).

Table 5.2. Overall occurrences of gesture types for cardholder and recipient

Gesture category	Role	Mean	Std. dev	Sum
Abstract descriptive	Cardholder	1.04	1.42	115
	Recipient	0.20	0.52	22
Iconic	Cardholder	2.88	3.10	320
	Recipient	0.59	1.34	65
Deictic	Cardholder	2.93	3.10	325
	Recipient	0.45	1.27	50
Pantomime	Cardholder	0.93	2.24	214
	Recipient	0.41	0.90	46
Metaphoric	Cardholder	0.43	0.78	48
	Recipient	0.07	0.29	8

Chi^2 comparisons of valence show no differences in the frequency of descriptive gesture for Positive or Negative experiences ($Chi^2_{(1)}$= 0.04, p = 0.83).

Given Cardholders make extra use of Abstract descriptive gesture a focussed comparison of frequency of those gestures alone was carried out but shows no difference with valence of experience ($Chi^2_{(1)}$= 1.08, p = 0.29).

5.3.2 Gesture duration for cardholder and recipient

To assess the duration of the different descriptive gestures a General Linear Mixed Models analysis (GLMM) was carried out on the average gesture duration. A GLMM with Valence (Positive vs. Negative) and Role (Cardholder vs. Recipient) as Fixed Factors and Participant as a Random Factor showed main effects of Role ($F_{(1,423)} = 78.6$, p <0.001) and Valence ($F_{(1,423)} = 78.6$, p <0.001) but no Valence × Role interaction ($F_{(1,423)} = 1.43$, p = 0.23).

The estimated marginal means show Cardholders' descriptive gestures lasted an average of 2.5 seconds whereas Recipient's lasted 1.7 seconds. The marginal means for the main effect of Valence show a smaller difference with gestures in relation to 'Negative' experiences lasting 2.3 seconds in contrast to 2.0 seconds on average for gestures to 'Positive' experiences.

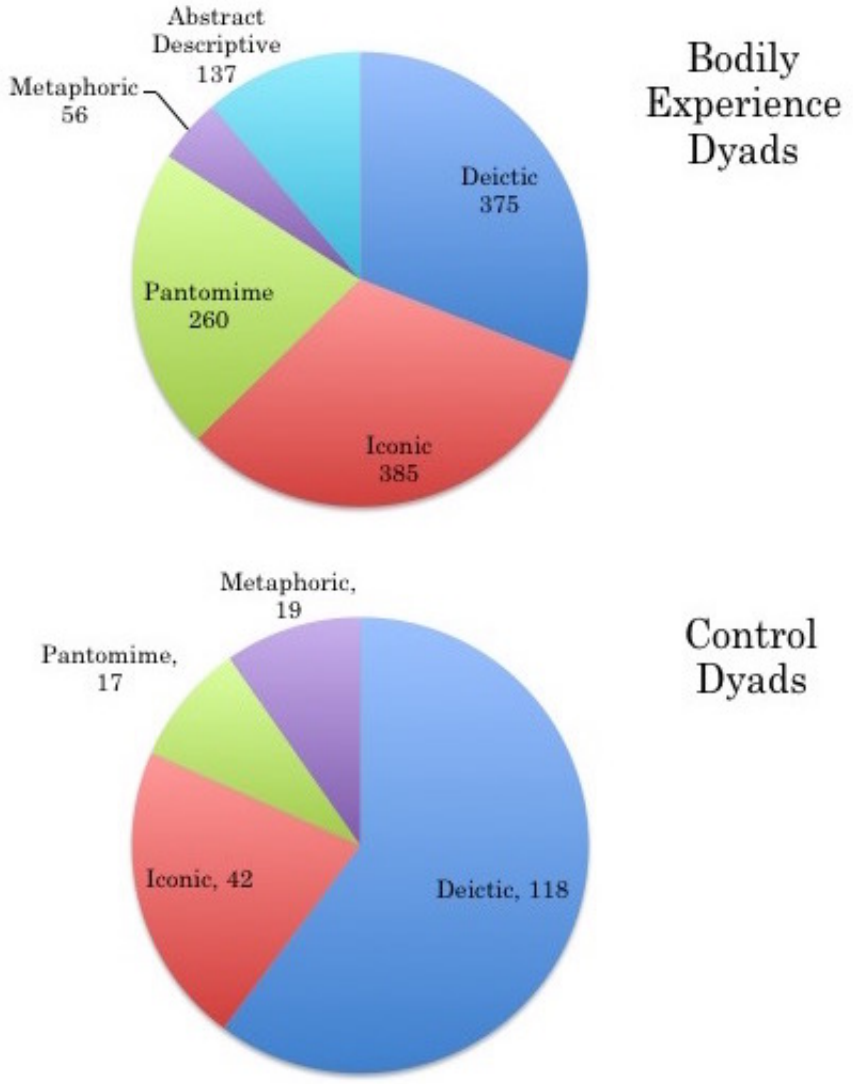
Because Recipients use very few Abstract descriptive gestures an additional GLMM analysis was carried out on the average duration of Iconic, Deictic, Pantomime and Metaphoric gestures alone. This shows main effects of Role ($F_{(1,350)}$ = 41.2, p <0.001) and Valence ($F_{(1,350)}$ = 31.7, p <0.001) as before but also a significant Valence × Role interaction ($F_{(1,350)}$ = 5.46, p = 0.02). Pairwise comparisons for this interaction show while both Cardholder and Recipient use gestures of longer duration for 'Negative' experiences the difference is more marked for the Recipient (+ 0.70 seconds) than the Cardholder (+ 0.29 seconds) ($F_{(1,350)}$ = 7.18, p = 0.01).

5.3.3 Comparison of gesture type distribution with control dyads

A full analysis of the control dyads in corpus 2 that is equivalent with the above analysis of the bodily experiences in corpus is not detailed here. This is because firstly; there not a compatible positive / negative valence contrast in the control dyads, as the interactions are about informal current affairs of a relatively neutral content, they are not easily separated into positive or negative experiences by virtue of their topic. Secondly; the structure of the interaction has no role equivalent as there is no cardholder or recipient, so there is no task related reason that participants would not contribute equally. For this reason an analysis of gesture distribution and length between interlocutors would not be directly comparable between the corpora. However, a useful comparison between corpora would be to analyse the distribution of descriptive gesture types to an insight into the distinguishing characteristics of how gesture are produced differently for conversations about bodily experiences in corpus 1 and more typical conversational topics in corpus 2.

Overall, the distribution of descriptive gesture type is different for the control dyads and the bodily experience dyads. One obvious difference is there are no Abstract descriptive gestures at all in the control dyads compared to 9% of the bodily experience dyads. The distribution of the remaining categories of descriptive gesture is also different ($Chi^2_{(3)}$= 61.6, p <0.001). As the Pie Charts in Figure 5.2 show, participants describing bodily experiences make proportionally more use of Iconic ($Chi^2_{(1)}$= 15.3, p <0.001) and Pantomime gestures ($Chi^2_{(1)}$= 17.4, p <0.001) and correspondingly less use of Deictic ($Chi^2_{(1)}$= 44.9, p <0.001) and Metaphoric gestures ($Chi^2_{(1)}$= 8.63, p <0.001).

Figure 5.2. Distribution of total descriptive gesture occurrences for bodily experience dyads and control dyads



5.4 Discussion

5.4.1 Effects of cardholder and recipient role on overall gesture frequency and duration

The task in corpus 1 ensued an interactional structure akin to a storytelling format. One participant adopts the teller role- describing a recalled bodily experience prompted by a topic written on a card. Their conversational partner adopts the recipient role as party to the

description, even so the recipient is still expected to interact with the cardholder/teller to show attention, comprehension or an attitudinal or emotional response to the description. The effect of role is marked on gesture frequency and duration, cardholders producing approximately five times as many descriptive gestures as recipients and their gestures last longer. This difference is expected as the cardholder's description takes precedence over the interaction- the cardholder holding the floor to perform more in-depth gesticulation.

It has been previously shown while listening passively to the cardholder's description recipients avoid gesticulating, producing little or no gesture during any backchannel feedback[1] (Gullberg, 1998). Recipients are found to gesticulate only when actively engaged to ensure comprehension- when producing more in-depth feedback responses, such as second stories (Heritage, 2011) or during clarification sequences (Healey et al., 2015). In the current analysis, this is reflected in Table 5.1 where cardholders produce almost five times the amount of descriptive gesture than recipients showing a prominence to the cardholder's gestural contributions. Further, an in-depth qualitative analysis in Chapter 8 looking directly at 6 excerpts from the corpus shows little gesticulation from the recipient over the cardholder's talk, in fact, care is taken by recipients to minimise overlap with deliberately placed verbal and non-verbal feedback around the cardholder's contributions.

Surely, when listening passively to a description it would difficult to be attentive if concurrently gesturing over another's speech[2]. Additionally, from an interactional standpoint, concurrent gesture over a cardholder's talk would be seen as a bid for the floor (Streeck and Hartge, 1992; Mondada, 2007) prompting a cardholder to subsequently stop and attend. Indeed, this would interrupt the flow of the interaction and interfere with the task.

5.4.2 Differences in gesture type distribution

The distribution of the descriptive gesture types is broadly the same for cardholders and recipients and across all the experiences being described with one exception; cardholders use a higher proportion of abstract descriptive gestures than recipients. This does not support the first pre-

[1] Observed situations where recipients use gesture during backchannel feedback are; motor mimicry (Bavelas et al., 1987) or affect burst expressions (Schröder, 2003)

[2] An exception to this would be the instructor/learner interactions in Furuyama (2000) but in these interactions attention is on the gesture itself with recipient active in a learning task

diction as the distribution of gesture type occurrences do not match between cardholders and recipients. If behavioural mimicry were a consequence of an automatic perception-behaviour link, it is expected the same type of gesture would be used by the cardholder and recipient. Yet, the above analysis shows abstract descriptive gesture is used differently in different roles.

By producing fewer abstract descriptive gesture, the recipient is seen to focus less on the sensation of the bodily experience but still engages with the contextual aspects of the experience. Returning to the issue Wittgenstein (1958) raises; a sensation is never truly shared as there is no direct access between person to person. An explanation could be the difference is down to an issue of politeness, perhaps expressions acknowledging private inner experience such as bodily sensation are seen to be intrusive by interlocutors. Heritage (2011) suggests interlocutors are sensitive to this asymmetry of definitive epistemic access:

> Because persons conceive experience as 'owned' by a subject-actor, and as owned in a singular way, a 'problem of experience' arises (Heritage, 2011).

Recipients may avoid mimicking the abstract descriptive gestures produced by the cardholder because do not hold epistemic rights to it. In an alternative explanation for gestural mimicry (see Section 3.2.7.1) interlocutors collaborate in the constructing shared gesture-speech units have meaning mapped onto their form- this establishes common gestural referents of meaning. These 'expressive referents' are then reused or adapted by one another to clarify and display understanding. However, considering the difficulty of knowing another's experience- it would be difficult to assign a gesture-speech unit to a privately inaccessible sensation by this type of abstract means. Results demonstrate iconic and deictic gestures were more frequent overall for both roles indicating interlocutors find it easier to refer to the cause of the sensation through iconic gesture or locate the sensation by pointing to it with deictic gesture. Using external situational context and/or concrete environmental details can be commonly shared to infer how the sensation is caused rather than directly attempting to describe a private experience.

5.4.3 Differing impact of valence on cardholder and recipient gesture production

Valence of experience does not impact on gesture type but does lead to longer gesture durations overall. When abstract descriptive gestures are excluded the results indicate this effect

is different for the different roles. This is incongruent with the second prediction, that cardholder's and recipient's gesture patterns would reflect any difference to one another due to valence. Recipients appear to respond *more* strongly to descriptions of negative experiences by extending the duration of the iconic, deictic, pantomime and metaphoric gestures they produce in response.

Some accounts report positive and negative valence has been seen to influence facial mimicry. Positive facial expressions are mimicked more than negative (Likowski et al., 2011), sad facial displays show inhibited facial mimicry (especially for out-group members / strangers) and counter-expressions to anger expressions (Bourgeois and Hess, 2008). It is suggested the perception-behaviour link is selectively moderated to induce facial mimicry if the expression serves an affiliative goal (Duffy and Chartrand, 2015).

The results reported here indicating the recipient engages more with negative experiences of pain could be said to be in contrast to the above account. Assuming that the positive experience descriptions would contain more happy expressions and the negative experience descriptions of pain would contain more sad expressions- congruent results would see both cardholder and recipient to mimic one another more for positive experience and mimicry would be inhibited for negative experiences. However, the analysis shows the opposite for both interlocutors who engage more gesturally for negative experiences- the recipient showing a larger increase.

Despite this, a description of negative experience could still enlist a 'moral obligation' to acknowledge and form empathetic communion as discussed above (Heritage, 2011), and for this reason still serve a affiliative goal. It could be suggested the results indicate this obligation is more influential when responding to descriptions of negative experiences as recipients engage more by extending the duration of the gestures they produce in response.

5.4.4 Gestures types useful for describing bodily experiences

The difference in the distribution of gesture type occurrence between the corpora demonstrating gesture is used selectively and resourcefully within interaction. Certain gesture types will be more useful to convey particular messages: such as deictic gesture to describe spatial or location information, as with a route description (Rieser, 2011; Mol et al., 2012).

For description of bodily experience interlocutors find concrete illustrative gestures (iconic,

deictic and pantomime) together with abstract descriptive gestures are particularly useful for describing bodily experiences. For example, when describing a toothache, the cardholder might add deictic gestures describing the location of the pain by pointing to it, or perhaps detailing they had to eat on the other side of their mouth, accompanied gesturally by pantomiming a chewing motion.

This indicates the task structure was successful in engaging interlocutors in communicating bodily experiences, supporting the anticipation the corpus contains empathetic interactions.

5.5 Summary

This chapter investigated the whether the two general models that propose a basis for automatic mimicry extend to descriptive or content-specific gestures observed in spontaneous interactions about bodily experiences. The first model, taken from research on facial mimicry (usually of photographic or video stimulus expressions), suggests that automatic imitation of facial expressions is the mechanism that simulatively underpins the experience of emotional contagion due to overlapping neural representations. When taken together with higher cognitive perspective taking leads to the experience of empathy (Hatfield et al., 1994; Hess and Blairy, 2001; Grèzes and Julie, 2006; Sinigaglia and Carrado, 2011; Gallese, 2013). The second model, taken from research on the perception-behaviour link, proposes that automatic behavioural mimicry of particular behaviours such as face-touch and foot-shaking with confederates is similarly explained by the perception of an action and the performance of an action sharing a common representational cognitive domain, but this is related to higher cognitive processes of perspective taking and not the affective dimension as the first model suggests (Chartrand and Bargh, 1999; Lakin et al., 2006; Chartrand and van Baaren, 2009; Chartrand and Lakin, 2013).

The analysis counted the incidence of descriptive gesture based on five categorical types over the duration of the interaction- (a measure similar to the studies that observe behavioural mimicry studies of foot-shaking / face-touching). This analysis also made a comparison of the length of descriptive gesture. Neither prediction based on these two models is directly supported by the results in a simple way; cardholders and recipients use different patterns of gesture type and the length of the gestures vary differently between gesture type and role. Notable is that participants behaviours are heavily influenced by their role (cardholder/recipient) as recipients

gesture many times less frequently and for shorter duration.

The main difference in gesture production between interlocutors was that recipients did not appear to mimic the abstract descriptive gestures performed by cardholders. In the research that states a perception-behaviour link as basis for behavioural mimicry, many moderators of mimicry are presented; all related to it's prosocial effect. So the mimicry of abstract descriptive could be explained as being inhibited by virtue of the recipient respecting the cardholder's epistemic access. Or rather, this could be related to higher cognitive processes that govern gesture production, as due to the inaccessible content of this type of gesture it would be difficult to establish a common gestural referent of meaning.

In this analysis the influence of valence on descriptive gesture is different between roles, with a longer engagement from recipients for negative experiences. This does not align with some studies of facial mimicry that find less mimicry in response sad expressions and mimicry does not occur for other negative expressions, such as anger (Bourgeois and Hess, 2008; Iacobini et al., 2009, 2010). In this instance, the more active engagement from the recipient in negative experiences described by the cardholder can be explained as influenced by an affiliation goal. So this difference is in line with the same prosocial causes that moderate mimicry in studies that propose a perception-behaviour link. Alternatively, it could be suggested that the negative experiences included in this corpus provide a repertoire of gestures that are longer in length.

Alternative to these two general models, there are several studies find mimicry of gestures relating to the content of the talk within interaction (see Section 3.2.7.1) but do not suggest simulative underpinnings. Rather these studies focus on a selection of examples showing how gestural mimicry is a strategically utilised display of expressive referents mutually and collaboratively constructed to establish shared concepts. It is proposed that the results point more solidly to that the cognitive processes that are used to produce descriptive gesture in these interactions are governed by higher systems that are interactionally involved and strategic, and not lower cognitive systems of automatic motor responses. This is particularly evident as the influence of topic and type of interaction evaluated in these corpora show a significant difference in the gesture types used.

Chapter 6

Posture congruence

6.1 Introduction

The previous chapter analysed the pattern of descriptive gesture occurring between interlocutors communicating about bodily experiences. The results were analysed with a view to determine whether simulative mechanisms of mimicry extended to gesture that is specific to the content of the talk. The results were not directly compatible with accounts describing mimicry as 'ubiquitous and engaged in automatically' (Hatfield et al., 1994; Hess and Blairy, 2001; Grèzes and Julie, 2006; Sinigaglia and Carrado, 2011; Gallese, 2013; Chartrand and Lakin, 2013) as cardholders and recipients are found to use different patterns of gesture type and length. Differences were found; firstly where recipients produced less abstract descriptive gestures in comparison to cardholders, and secondly, cardholders produced longer gestures overall in response to descriptions of negative over positive experiences whereas cardholder showed no difference. It is suggested that descriptive gesture is not governed by lower cognitive systems of automatic motor responses but higher cognitive systems that are conscious, strategic and interactionally involved.

A limitation of this analysis is the coarseness of the comparison- based on the assumption if a gesture is mimicked, both the imitation and the original will be categorised as the same broad gesture type, this leaves only the potential that there is an incidence of mimicry. Therefore, as the global distribution of gesture types was shown to differ between cardholder and recipient, there is still the possibility of behavioural mimicry to be present within the instances of gesture

that did occur, albeit not as uniformly across all gestural expressions an automatic perception-behaviour link would suggest. The analysis in this chapter uses a more direct approach to detect similarity between interlocutors expressions by comparing their morphological form within their overall posture but also measures overall posture congruence.

Body movement data makes it possible to directly, quantitatively test the degree of similarity between body postures found in dialogue. A method set out by Yang et al. (2010) to compare postures in motion capture data provides a similarity value for the angle between corresponding joints (see Section 4.5.3). Although this approach gives less insight into the content-specific characteristics of these expressions, as the comparison of gesture type in the previous chapter does, the measure will determine whether interlocutors bodily expressions are similar in form (as opposed to type) over the course of an interaction. As with the previous analysis, the current analysis places no time restriction on what is considered mimicry, as recurrence of similar body posure sequences are considered over the entire interaction.

The prevalence of posture congruence has been widely observed as well as positively correlated with rapport and shared viewpoints (see Section 3.2.1), it is regarded as an example of nonconscious behavioural mimicry. Section 3.2.6 details studies suggesting others' affective states and attitudes can be identified from their outward manifestation in body postures. It is suggested that adopting a similar posture when interacting communicate messages indicating an understanding of others' affective states and attitudes, increasing sense of involvedness, togetherness, and consequently, rapport. As with behavioural mimicry, it is suggested that this tendency to adopt similar postures is engaged in nonconsciously and automatically (Chartrand and Lakin, 2013). This analysis tests the validity of this hypothesis by measuring postural congruence in a corpus of affect-laden descriptions of bodily experience.

A limitation for the majority of studies observing postural congruence, is postures are compared against a set of predefined body configurations - for example left arm down or left arm up and so on. Observed postures are classed (manually coded or computationally grouped) as a particular predefined body configuration by virtue of being more similar to it than to other body configurations in the set. It is then the identified classes the postures are judged similar to that are then compared to determine a binary measure of posture congruence. In addition, postures are classified at intervals, sometimes minutes apart, disregarding any postural activity between. These studies focussed on select body parts, such as just head and torso, or head

and arms. When these experimental factors are considered together it could be argued these studies paint a lossy and approximated picture of postural activity.

The method used below side-steps these issues by allowing for a granular, full-body and non-binary measure of postural similarity over a continuous time-series. Detailing how much similarity is present between interlocutors posture's at any given time. The approach can also account for posture similarity occuring at an offset- for example a copied posture may not appear simultaneously to the posture it mimics but later in the dialogue, by detecting sequences of recurrent similarity over all time-offsets.

The following analysis uses the measures of posture similarity as set out in Section 4.5.3.3. Data from a participant's skeleton (containing rotation angles for each joint) is compared to another participant's skeleton to calculate similarity values for a pair. This produces a matrix of values; for every time unit (10ms) a comparison is made to every time offset for that pair. For example, skeleton data for the first 10ms for the first participant is compared to skeleton data of the second participant at 10ms and all other time units in that item. The measures used in this analysis are as follows:

1. Overall average recurrent posture similarity. The similarity values considered recurrent are over a threshold of the mean similarity for that pair.

2. % Determinism. The average proportion of recurrent posture sequences over a set time threshold. The time threshold is set as the average length of a gesture as hand annotated.

3. MAXLINE. The longest consecutively recurrent diagonal sequence of posture similarity between the pair.

As covered in Section 4.5.3.3 there are three constraints on posture similarity considered in this analysis. Firstly, each person has individual physiological bodily constraints on how they can move. Each person also has characteristic movements that form mannerisms unique to them. These factors will vary from person to person. The Self-Identity data set compares an individual participants similarity values with themselves at all offsets forming the matrix of similarity values described above. Measures from this data set will reveal how much a persons posture can vary despite these constraints. As this analysis is taken from a corpus of dialogues resembling narrative structure; whereby one participants describes a bodily experience and their partner is recipient, there will be an effect of role on the dialogue.

To measure the effect of role on similarity the Self-Min data sets takes a participant in a cardholder role and compares it with themselves in a recipient role for the same topic. These data sets are used as baselines to account for factors constraining posture.

A third data set pairs random cardholders and recipients of the same topic to determine how much similarity occurs purely by chance, this is compared with the data set of real interactional pairs. The difference between real and randomly paired dyads will show how much interacting partners posture is similar over chance levels.

Table 6.1. Four discrete data sets (for further description see Section 4.5.3.3)

Data Set Name	Description	Compares	Role	Topic
Self-Identity	Self similarity data indexes the degree in which a single participants posture varies over the course of a single dialogue	self/self	same role	same topic
Self-Min	Self similarity data in reciprocal role indexes the degree in which posture varies between describing bodily experiences and responding on the same Topic with the same subject	self/self	different role	same topic
Chance	Randomly matched pairs similarity data indexes the degree in which participant's posture is similar by chance	self/other	different role	same topic
Real	Similarity data for real interactions of actual pairs	self/other	different role	same topic

6.2 Predictions

1. Automatic posture matching: if interlocutors tend to automatically match one-another's posture during conversation their average overall posture similarity is expected to be reliably greater than chance.

2. Strategic posture matching: if posture congruence is used selectively in support of specific communicative goals posture similarity is expected to show an increase at specific points during the interaction.

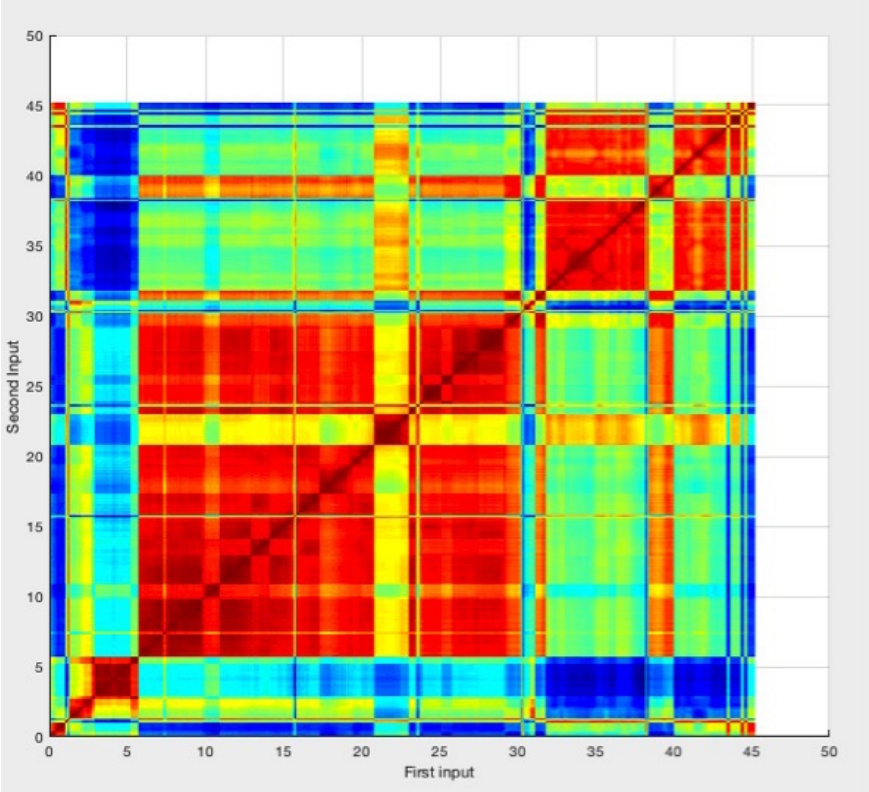
3. Autonomous posture: if interlocutor's body posture is influenced solely by the organisation of receiving a message or delivering one; where postures are adopted independently of a partner's posture, there will be no indication posture similarity is above chance levels.

6.3 Results

6.3.1 Self-similarity

Figure 6.1 shows an example of a recurrence plot displaying self-similarity for the Self-Identity data set. There is always a diagonal line of maximal similarity corresponding to self-identical posture at zero offset. These plots also commonly feature a chequered pattern of symmetrical blocks across the whole plot corresponding to people returning to self similar postures throughout the conversation; demonstrating people have a tendency to stand or gesture in individually characteristic ways.

Figure 6.1. Self-identity similarity CRQ plot



The self similarity data can also be used to assess how much posture is effected by virtue of taking a Cardholder or Recipient Role, and the influence of the Topic of the experience being described. A GLMM linear analysis of recurrent posture similarity over the Self-Identity data set was carried out with Role (Cardholder vs Recipient) and Topic (Backache / Laugh / Yawn / Toothache); their interaction set as Fixed Factors and Participant as a Random Factor. This shows an effect of Role ($F_{(1,136)} = 50.4, p < 0.001$) and an effect of Topic ($F_{(3,136)} = 4.12, p = 0.01$) but no interaction between Role and Topic ($F_{(3,136)} = 0.14, p = 0.94$). This demonstrates, for example, posture adopted whilst being Recipient to a description of a Yawn is not significantly different to the posture adopted whilst being Recipient to a description of a Backache.

The effect of Role shows the postures adopted when describing a bodily experience (Cardholder) are consistently less self-similar (mean = 0.964) than when they are Recipient (mean = 0.976). The analyses in Chapter 5 Section 5.4.1, found Cardholders produce 84% of overall descriptive gestures; indicating Cardholders move more to produce gestures to help illustrate what they are

saying. This is to be expected as by moving more Cardholders will necessarily vary their posture more, such variations in posture will entail less self-similarity at time offsets where postures do not match. As Chapter 5 details it has been shown when listening passively Recipients avoid gesticulating (Gullberg, 1998; Battersby, 2011; Healey et al., 2015), resulting in little variation in posture. By adopting a similar posture throughout a Recipient's posture will be more self-similar at different time offsets.

Pairwise comparisons for Topic indicate that taken together the Cardholder and Recipient tend to vary their respective posture least when talking about Laughing and most when talking about a Yawn and a Toothache (Yawn - Laugh: $t_{(136)} = -3.15, p > 0.001$, Toothache - Laugh $t_{(136)} = -2.47, p = 0.02.$). All other comparisons were not significant. Table 6.2 illustrates the pattern of self similarity of posture for the four different topics.

This pattern corresponds to cursory observations of interlocutors tending to raise their hands to point to their mouth when talking about toothaches and raise their arms to demonstrate stretching outwards when talking about yawning; both stereotypical actions for these topics cause participants to vary their posture from the usual stance of having their arms by their sides.

Table 6.2. Estimated marginal means for overall posture self-similarity

Topic	Cardholder	Recipient
Yawn	0.961	0.973
Toothache	0.962	0.975
Laugh	0.967	0.978
Backache	0.965	0.976

The Self-Min data set compares the same Participant in different roles for the same Topic. Comparing self similarity but in a different dialogue. This indicates the degree of self similarity of Participants to themselves when they are in a Cardholder Role compared to a Recipient Role. Comparison of Self-Identity and Self-Min thus provides an estimate of how much variations in a person's posture are due to being in the same Role (Cardholder or Recipient) in the same dialogue and different Roles (Cardholder vs. Recipient) in different dialogues. A simple GLMM linear analysis with Data Set as a single Fixed Factor and Participant as Random Factor shows

a main effect of Data Set ($F_{(1,210)} = 9.77, p < 0.001$). The overall mean self-similarity for Self-Identity is 0.967 for Self-Min 0.963. This shows only a small decrease (comparatively to Other-similarity detailed below) in overall posture similarity when compared to themselves in different Roles and different dialogues, indicating people do not tend to repeat their own movements within a dialogue and do not use different expressions when taking a Cardholder or Recipient Role.

6.3.1.1 Determinism

Determinism gives a percentage of recurrent sequences over a time threshold of the average length of a gesture. Determinism for the Self-Identity data set measures the degree to which peoples repeated movements occur in sequences. For Self-Identity this measure is positively skewed and therefore analysed with a GLMM Gamma regression with Role (Cardholder vs Recipient) and Topic (Backache / Laugh / Yawn / Toothache) and their interaction as Fixed Factors and Participant as a Random Factor. This a reliable effect of Role ($F_{(1,136)} = 75.6, p < 0.001$), no main effect of Topic ($F_{(3,136)} = 0.52, p = 0.67$) and a marginal interaction ($F_{(3,136)} = 2.71, p = 0.05$). Overall people's body postures are much less deterministic as Cardholder (57%) than as Recipient (76%). Recipient's self determinism is highest level of recurrence found in all data sets. This suggests the posture people adopt when standing listening to someone else is the most common recurrent body configuration observed in this task.

Although there is no main effect of Topic, pairwise comparisons show the marginal interaction is due to Cardholder's posture for Toothaches being less deterministic (53%) than their posture for Yawns (61%) ($t_{(136)} = 7.84, p = 0.05$) whereas Recipient's posture shows no reliable difference. All other Topic $\times$ Role pairwise comparisons were not significant.

As previously seen above, out of the four Topics interlocutors tend to vary their overall average posture most when talking about Yawns and Toothaches. This suggests although the posture used by Cardholders to illustrate Yawns and Toothaches entail more variation of posture; the typical posture sequences associated with describing Yawns are more consistent and sustained leading to a more determinism in posture, and the typical posture sequences associated with describing Toothaches are shorter in length and more sporadic in form leading to less determinism in posture. For example, the standard movement to illustrate a yawn could be described

as lifting arms up to shoulder height and stretching out. However, a standard movement for illustrating toothaches out could be described as lifting the hand towards the mouth and pointing to the various locations of a toothache, a movement that is more fleeting and has more variation in the location pointed at. These cursory observations of the video data of the corpus suggest some indication that, amongst the topics used in this task, Yawns may generate more stereotypical postures; i.e. when describing Yawns Cardholders produce more conventional gestures associated with that particular Topic. The means reported in Table 6.3 show the overall pattern.

Table 6.3. Estimated marginal means for determinism in posture self-similarity

Topic	Cardholder	Recipient
Yawn	61%	73%
Toothache	53%	80%
Laugh	54%	74%
Backache	60%	76%

Comparison of determinism for Self-Identity and Self-Min provides an estimate of how consistent a person's posture is in the same role in the same dialogue and between different Roles in different dialogues. A simple GLMM linear analysis with Data Set as a single Fixed Factor and Participant as Fandom Factor shows no main effect of Data Set ($F_{(1,210)} = 1.83, p < 0.18$). Suggesting repeated sequences do not occur more over the same dialogue than a different dialogue in the opposite role.

6.3.2 Other-similarity

Figure 6.2 shows an example of a recurrence plot of an actual pair- displaying similarity data when a participant is compared to their interactional partner. These tend to be dominated more by horizontal bars. This pattern arises when one person's posture changes while the other's remains relatively constant. For example, if the Cardholder's and Recipient's postures match in the first five seconds of the conversation and then the Cardholder moves but the Recipient remains relatively static then the Recipient's posture at all subsequent times will remain similar to the Cardholder's initial posture. Posture matching appears to happen at times when the

Cardholder adopts the Recipient's passive pose. This is reflected in the recurrence plot as a horizontal block.

Figure 6.2. Other similarity CRQ plot

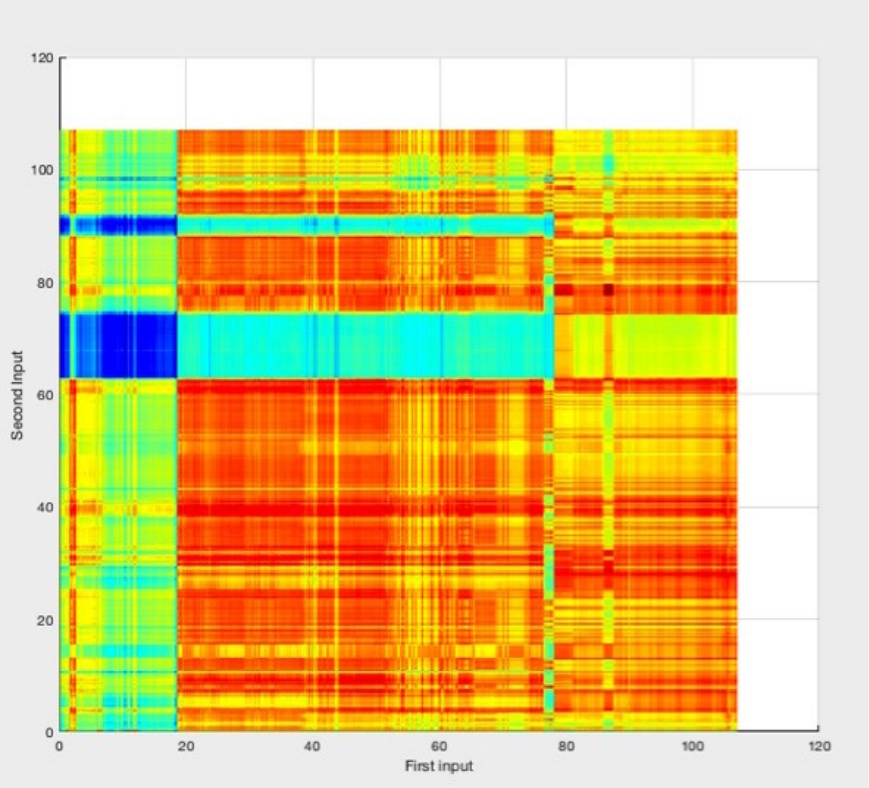
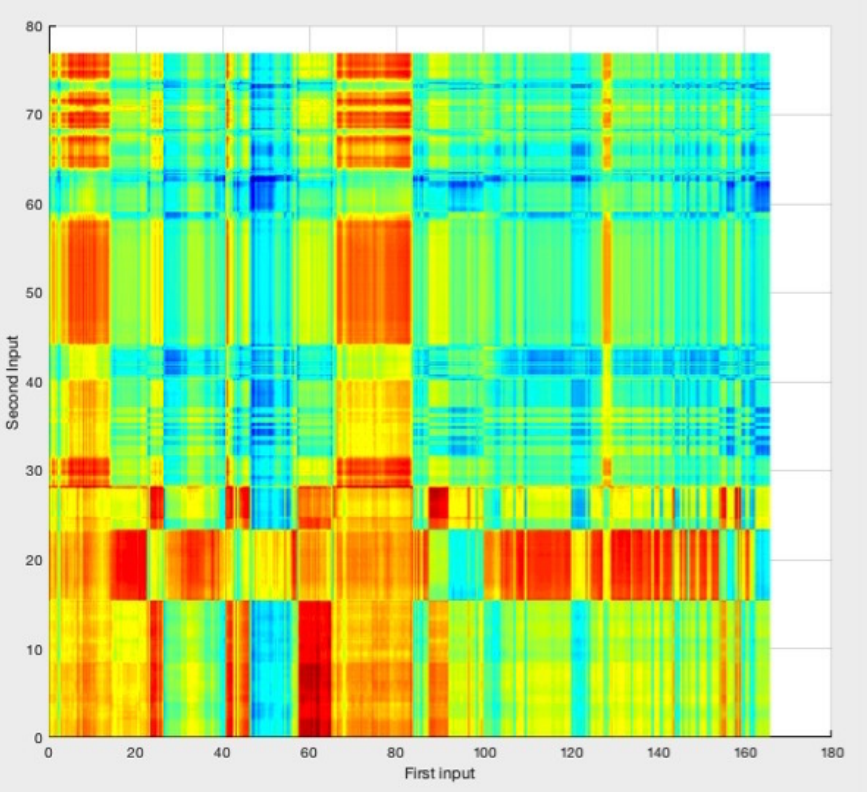


Figure 6.3 shows an example of a recurrence plot of randomly paired participants- displaying chance-similarity. These tend to show a more 'speckled' pattern of smaller horizontal and vertical bars which suggests randomly paired participants do occasionally match one-another's posture by chance but these matches are shorter and more randomly distributed.

Figure 6.4 shows the average posture similarities for self and other-similarity. In all cases these are all relatively high varying between 0.93 and 0.97. As would be expected, people are consistently more self-similar even over a range of temporal offsets (SelfID) and across dialogues (Self) than they are to other participants. This is due to the fact people impose individual physiological constraints on their range of possible movements, and tend to produce idiosyncratic movements and mannerisms. The figure shows average posture similarity to interactional partners is close to chance levels. This is not compatible with prediction one suggesting for interlocutors to automatically match one-another's posture their average overall posture similarity is expected to be reliably greater than chance.

Figure 6.3. Chance similarity CRQ plot



A GLMM analysis of other-similarity was performed with Data Set (Real vs. Chance) and Topic (Backache / Laugh / Yawn / Toothache) as Fixed Factors and Participant A × Participant B as a Random Factor. This shows no main effect of Data Set ($F_{(1,136)} = 0.95, p = 0.33$), a main effect of Topic ($F_{(3,136)} = 4.34, p = 0.01$) and no Data Set × Topic interaction ($F_{(3,136)} = 0.39, p = 0.76$). This shows recurrent similarity is not significantly different between real and chance pairings regardless of Topic.

Pairwise comparisons for Topic show recurrent similarity is highest for Backache and lowest for Toothaches and Yawns (Backache - Toothache: $t_{(136)} = 3.36, p < 0.001$; Backache - Yawn $t_{(136)} = 2.52, p = 0.01$; All other pairwise comparisons were not significant). This is also evident in chance pairings as seen in Table 6.4; showing the marginal mean levels of recurrent similarity for each Topic in the Real and Chance pairs. Chance values follow the same pattern of similarity as Real pairs, with lowest similarity for Toothaches and Yawns. This is in line with self-similarity data displaying the lowest similarity values for Toothaches and Yawns due to Cardholders varying their posture more for these Topics.

Figure 6.4. Mean recurrent posture similarity for each data set (see Table 6.1)

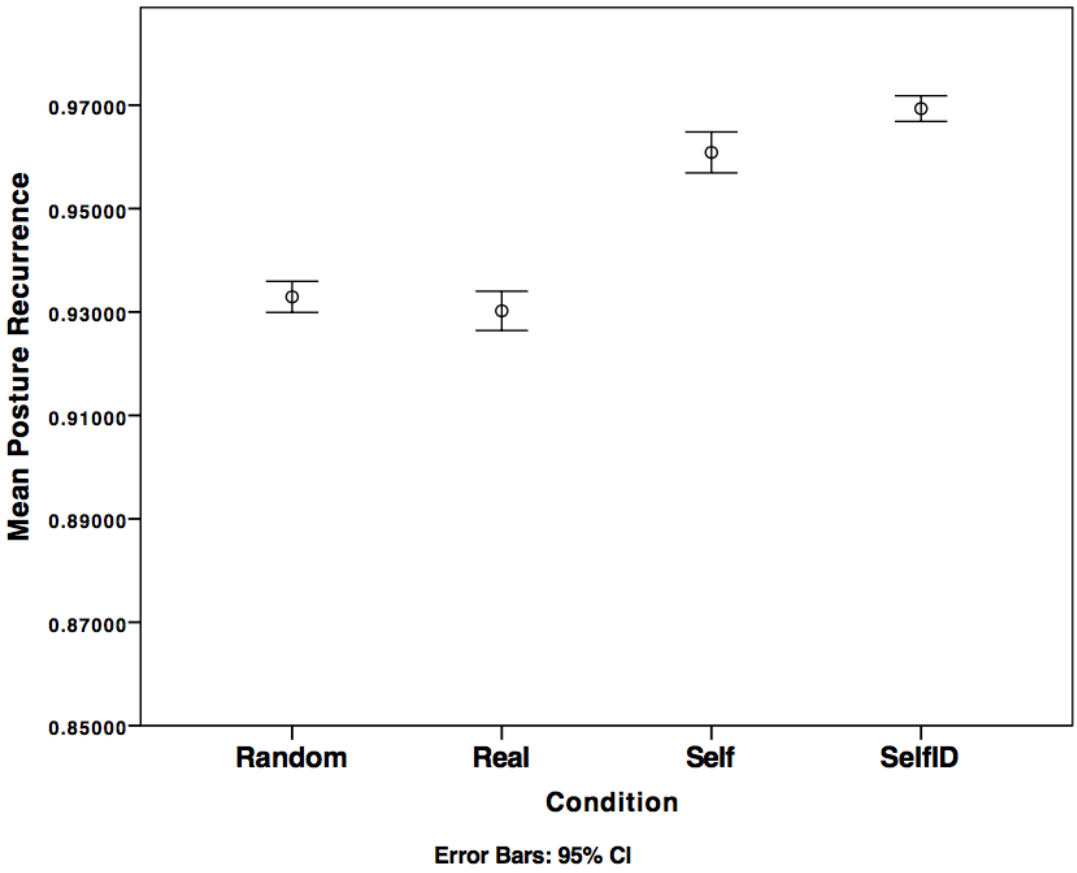


Table 6.4. Estimated marginal means for posture other-similarity

	Real	Chance
Yawn	0.929	0.932
Toothache	0.926	0.927
Laugh	0.931	0.933
Backache	0.933	0.939

6.3.2.1 Determinism

A GLMM analysis of determinism with Data Set (Real vs Chance) and Topic (Backache / Laugh / Yawn / Toothache) and their interaction as Fixed Factors and the two-way interaction between the two participants as a Random Factor shows no main effect of Data Set ($F_{(1,136)} = 0.82, p = 0.37$), no main effect of Topic ($F_{(3,136)} = 1.04, p = 0.38$) and no reliable interaction ($F_{(3,136)} = 0.13, p = 0.94$). Planned pairwise contrasts of Real vs Chance determinism for each Topic are not significant. It should be noted postures considered recurrent are based on a deliberately generous criterion designed to ensure even marginal cases of posture similarity are included. Regardless, Real pairs have no more repeated sequences than would occur by chance.

6.3.2.2 MAXLINE

MAXLINE measures the longest single recurrent sequence, i.e. the longest recurring posture sequence for a pair. A sequence is considered recurrent if the posture similarity value consecutively falls over the average similarity for that pair. This measure is positively skewed so a gamma distribution is used for analysis. A GLMM analysis with Data Set (Real vs Chance) and Topic (Backache / Laugh / Yawn / Toothache) and their interaction as Fixed Factors and pairs of participants as a Random Factor shows no main effect of Data Set ($F_{(1,136)} = 0.02, p = 0.90$), no main effect of Topic ($F_{(3,136)} = 1.49, p = 0.23$) and no reliable interaction ($F_{(3,136)} = 0.74, p = 0.53$). This shows the longest recurrent sequences are not significantly different between Random and Real pairs, nor are they different depending on Topic.

6.4 Discussion

There is no evidence from this analysis supporting prediction one, that people tend to automatically adopt similar postures during interaction. Real pairs show no more overall average posture similarity and no more repeated sequences than would occur by chance. Instead the more a cardholder moves and gesticulates, the less similar people's posture is- both to themselves and to their conversational partners over time.

There is also no specific evidence in line with prediction two; that posture congruence is used selectively to support communicative goals. However, instances of postures similar for strategic communicative purposes may not show up in these measures if they are infrequent as posture similarity is compared from overall recurrence averaged over an entire interaction.

The results are somewhat in line with prediction three, that interlocutors postures are adopted independently of a partner's posture but are influenced by the organisation of communication itself. Or more specifically; according to their conversational role and according the topic of conversation. This is seen in the data as a recipient's self determinism has the highest level of recurrence found in all data sets; indicating that recipients do not vary postures as often cardholders because they are relatively still while they passively listen. Also, topics that involve the cardholder moving more (Toothaches and Yawns) entail interlocutors postures are less similar than other topics; demonstrating the influence a topic has on posture.

Two additional components that differentiate people's posture from one another are differing physiological constrains and individual habitual patterns of movement. These constrains influence the range of postures people customarily adopt during conversation, allowing people their very own *style* of posture. However, it is still the case that individual postural patterns of movement are modified according interactional structures as above.

These results are in contrast to findings of posture similarity in LaFrance (1979, 1982), suggesting that interlocutors hold posture congruence to show attention and comprehension, as there would be a level of posture congruence above chance. As this analysis uses a full bodied and highly sampled method to measure posture similarity there is the opportunity to detect ongoing movement patterns that may cause divergent postures. The broad measure comparing discrete posture configurations used by LaFrance (1979, 1982) may overlook these divergent postures leading to a false positive result.

These findings are not directly incompatible with Shockley et al. (2003); indicating the postural sway of a pair converge when interacting as postural dynamics entrain due to vocal patterns and breathing rhythms related to speech delivery. However, this could be due to the full bodied analysis used here throws up differences Shockley et al. (2003)'s experiment did not measure. Only two sensors were used to detect postural sway, positioned on subjects torso and head. As seen in the current results movements that include moving hands to the face or stretching

out the arms contributed less posture similarity- suggesting arms and hands have a significant influence on postural similarity. It could be suggested by omitting the arms and hands there could by an underlying entrainment of torso and head movements.

6.5 Summary

Although the prevalence of posture congruence has been widely observed in previous studies (Scheflen, 1964, 1972; LaFrance, 1979; Trout and Rosenfeld, 1980; Maurer and Tindall, 1983; Feese et al., 2011; Tia et al., 2011; Hagad et al., 2011), a full bodied and highly sampled analysis of posture similarity found no evidence of postural congruence. The hypothesis that there is a nonconscious and automatic tendency to adopt similar postures, as with behavioural mimicry, is not supported by this study. The results for overall other-similarity indicate the average posture similarity between interlocutors in actual conversations is not reliably different from that which would be expected by chance. This study does not directly challenge the notion affective states and attitudes can be identified from an outward manifestation in body postures, however the notion that postural congruence communicate messages indicating an understanding of these states is not observed in this corpus of affect-laden descriptions of bodily experience.

Observed in the present analysis is that participants adopt postures that designed to assist in the act of interaction itself. The results for overall self-similarity show people vary their posture more when speaking than as recipient. This is due to cardholders taking the lead in articulating, verbally and non-verbally, their recalled experience. The results for determinism confirm this picture suggesting people hold the same posture longer as recipient than as cardholder, this is in line with research observing that a recipient is still when someone is speaking to them in order to be able to attend to what they are saying (Gullberg, 1998; Battersby, 2011). This shows a significant effect of the narrative dialogue structure used in this task on the postures adopted by participants in the cardholder and recipient roles, as opposed to an effect of automatic postural matching.

The results do show people tend to use certain characteristic movements to describe particular bodily experiences. This is shown in the self-similarity data- where cardholders vary their posture more when talking about toothaches and yawns, which often involve pointing to a tooth when describing toothache or stretching out when describing yawns. Cardholders move

less when talking about backaches and laughs where the hands tend to be kept lower. This result is also reflected in the other-similarity data where cardholders are less similar to recipients when talking about toothaches and yawns, but as demonstrated in the self-similarity data, this is simply by virtue of producing more dynamic postural activity for these particular topics and does not appear to be influenced by the interaction.

Also, this analysis shows physical constitution and habitual patterns of movement have a significant influence differentiating interlocutors posture. Posture appears to be autonomous, or directed by constraints related to the act of communicating rather than automatically matched to an interactional partner.

A limitation of this analysis is that there is not enough evidence to dispute that there could still exist infrequent instances of selective postural congruence in these conversations, as the analysis is based on posture similarity averaged over an entire interaction. Also, this analysis only covers full body postural comparisons so any partial postural congruence is overlooked.

Part IV

Interactional strategies toward intersubjectivity

Chapter 7

Responding to embodied experience

7.1 Introduction

Understanding is the ability to act appropriately (Wittgenstein, 1958, p. 58).

As the results in previous analyses did not observe any clear evidence of automatic behavioural mimicry in interlocutors' gestures or postures when talking about bodily experiences, this leaves the question of what do interlocutors do when empathetically engaging with one another's personal accounts. In what ways is mutual embodiment significant to how people work toward intersubjective understanding?

In order to answer this question a closer analysis of how people nonverbally respond to one another's talk is required. Specifically, the interactional context of feedback behaviours are investigated with a view to identify ways in which a recipient acknowledges or engages empathetically with a cardholder's descriptions of bodily experience; and how a cardholder acknowledges or engages with a recipient's response, with a particular emphasis on nonverbal feedback.

Feedback behaviours, as described in Section 3.2.8, are verbal or nonverbal responses to an interactional partners contributions. Their function ranges from showing attention, claiming or / and exhibiting understanding or uncertainties in understanding, to producing an attitudinal or emotional response to the message a cardholder is trying to convey.

Showing affiliation to an account of a personal experience has been found important to the structure of interactions about personal experience; by affirming a cardholder's stance as legitimate and understood a recipient encourages the ongoing telling activity Kupetz (2014). Heritage (2011) describes a series of empathetic responses holding different degrees in which they embody a cardholder's experience to show affiliation.

1. The least empathetic response Heritage describes are ancillary questions; these present a question related to the description, but forgo empathetic engagement with the experience as they shift the focus of the conversation away from the experience.

2. Parallel assessments; These describe a similar experience supporting or 'seconding' the cardholder's description. These assessment show a personal understanding but decline directly engaging with the experience of the cardholder.

3. In contrast, subjunctive assessments do attempt to directly engage with the cardholder's experience by saying if they were to have the same experience they would feel similarly.

4. Claiming imaginary access to the experience as would-be onlookers or to the experience described by cardholders.

5. Heritage suggests the most empathetic response is what Goffman (1978) describes as response cries; is an exclamatory interjection defined as situational act displaying communicative alignment with on-going or told events such as expressions of revulsion, strain, pain, glee or surprise. Response cries convey a strong sense in which a recipient is affected by the description. These short, mostly prosodic sentiments register affiliation without interrupting the telling activity but set up a stance to later give more propositional contributions of understanding and affiliation. Schröder (2003) terms these 'affect bursts' (see Section 3.2.8); short prosodic utterances and facial expressions displaying a strong emotional or attitudinal reaction to an experience. Similarly, Bavelas et al. (1987) classify empathetic recipient responses such as these as motor mimicry (see Section 3.2.7.2). These short displays show an understanding of the type of experience described and how that experience affected the cardholder.

Kupetz (2014) discusses the form these resources take within an affect-laden dialogue. Observing the sequence of empathetic responses starts from fleeting and nonverbal from the beginning of a description. This incorporates facial expressions/expressive movements, facial expressions

alongside response cries and / or short assessments (such as "That's awful!"). These would fall under Heritage's last and most empathetic response as detailed above- including 'affect bursts' and 'motor mimicry'.

Mid-sequence Kupetz (2014) observes ancillary questions and assessments using mental words (such as "I think ... "). These are akin to the three feedback resources least empathetic according to Heritage (ancillary questions, parallel then subjunctive assessments). Additionally, Stivers (2008) shows a mid-sequence nod positioned where the speaker offers a telling element providing insight into their stance displays affiliation with the speaker's position. Sequential assessments would be too strong at this point, however nods can be treated as affiliation tokens endorsing and accept the storyteller's perspective at this stage of the description. Nods show alignment with the cardholder's stance while respecting the asymmetrical activity of storytelling, allowing the cardholder to continue. In this sense, the recipient's nod can be treated as an empathetic feedback (Stivers, 2008).

The most substantive and verbose empathetic responses tend to occur later in a dialogue of descriptions of personal experience, such as second stories detailing similar experiences or formulations summarising or paraphrasing the experience back to the teller. Although more substantive and comprehensive; the latter responses are not deemed as empathetic as the fleeting nonverbal responses tending to occur in the earlier stages of a description.

A limitation to Kupetz' and Heritages' observations is they are largely focussed on the verbal aspects of empathetic responses. How nonverbal contributions to these resources are outlined above are analysed in this chapter.

At times, people use feedback to signal uncertainties in their understanding of another's talk (Clark and Schaefer, 1989). Here, this is defined as a type of repair: when a speaker recognises an error in speech, or their recipient does not fully comprehend what was said, and works towards resolving the problem by repeating what has been said, correcting or clarifying their talk (see Section 4.7). Moments of repair are regarded as critical for intersubjectivity, Healey (2008) argues resolving moments of repair are more important for the co-ordination of understanding than positive evidence of understanding as interlocutors actively collaborate towards mutual understanding. Feedback in clarification sequences represents an explicit form of these moments.

This analysis looks at the embodied resources used during the process of repair, such as gesture (Gullberg, 1998; Holler and Wilkin, 2011a). It has been shown recipients increase the frequency of their head nods when a speaker exhibits disfluences (self-repair) in their speech in order to help resolve the problem (Healey et al., 2013), this type of nonverbal feedback is particularly useful to recipients in this context as it does not interrupt the speaker's turn. Embodied resources are also useful at these points as if an interactional resource is not suited or insufficient interlocutors are expected to turn to alternative or multi-modal resources to compensate, designing their contributions with their interactional partner in mind. This analysis extends this line of inquiry by looking at how interlocutors use gestural resources during clarification sequences asking how embodied feedback is valuable to interlocutors when problems in mutual understanding occur.

Two task roles are distinguished as per the previous analysis: 'Cardholder' for the participant currently relating their experience on their card and 'Recipient' for the participant who is audience for the description. An additional distinction for analysis of clarification questions defines two dialogue roles for clarification sequences: 'Speaker' for the participant currently talking and who's turn it is and 'Non-speaking addressee' for the participant not currently speaking. Although the cardholder is typically the speaker in this task- recipients still provide verbal comments and feedback. On these occasions the recipient takes the speaker role and the cardholder becomes the non-speaking addressee. For this analysis turns are coded from a transcription of the dialogue. The participant producing an utterance is coded as the speaker, including utterances spoken overlapping with an utterance spoken by the other participant.

For analysis of experience type, target items were separated intuitively into either a 'Positive' valence category (Satisfying yawn, Laughing out loud, Back massage) or a 'Negative' valence category (Backache, Stomach ache, Toothache). Feedback behaviours have been separated into three function types: Contact / Perception (C/P), Comprehension (C) and Attitudinal / Emotional (A/E). Each participant's feedback contributions in each dialogue is hand coded, specifying whether feedback occurs, whether it is verbal or nonverbal and which function type it falls under for both corpora (see Section 3.2.8).

Descriptive gestures are coded as per previous analyses- pantomime, metaphoric, deictic, abstract descriptive and iconic (see Section 4.6.5.3). Additionally, motion analysis provides an index of 'significant hand movement'- a threshold of any acceleration value of the fastest moving

hand one standard deviation from the mean movement for that person is used to determine for what is considered significant motion, any movement above this threshold is considered significant (see Section 4.5.2 for a full description of this method).

Clarification sequences are additionally coded into two stages: the question itself- what is asked when someone fails to fully comprehend something in another participants previous utterance (coded as CQ), and the response to the clarification question (coded as R), see Transcript 7.1 and Transcript 7.2 for examples of this coding. Responses to clarification questions are not always verbal - they can be in the form of a nod or point towards the listener to confirm they are correct (described as interactive deictic gesture in Section 4.6.5.3).

```
1. CH: and then you know I'm reaching back like this
        (0.1)
2. NCH: is it like a slipping (disc type thing?) (CQ)
3: CH: (this thing) I mean no it's like a muscle I think got caught or knotted up (R)
```

Transcript 7.1. Clarification sequence: Example 1

```
1. C: a couple of books about that thick
        (0.4)
2. D: how thick?= (CQ)
3: C: =er two inches thick or something= (R)
4: D: =ok
```

Transcript 7.2. Clarification sequence: Example 2

Drawing from Corpus 1 (bodily experience dyads) as laid out in Section 4.2 comprising of a collection of dyadic conversations about six different bodily experiences, this section presents an analysis of feedback behaviours and clarification sequences produced by the pairs of interlocutors for each experience type.

The distribution of the different types of verbal and nonverbal feedback signals found in Corpus 1 are then compared with Corpus 2 (control dyads) as laid out in Section 4.3, comprising of a collections of dyadic conversations about current popular topics. A comparison across both corpora give an insight into the distinguishing characteristics of how feedback differs for conversations about bodily experiences in corpus 1 and more typical conversational topics in

corpus 2. As the control dyads did not have the same role distinction as the bodily experience dyads a comparison analysis of clarification sequences is not useful as there is no direct comparison.

7.2 Predictions

Overall recipients and cardholders are hypothesised to organise their feedback responses appropriately to descriptions of personal experiences, the pattern of this organisation will differ from the organisation of feedback produced in interactions of other types of dialogues. Heritage and Kupetz' resources discussed above create specific expectations as to the pattern of feedback types generally produced when talking about bodily experiences, these are outlined below:

Response cries / affect bursts / motor mimicry- This type of feedback would fall under nonverbal A/E for the facial expressions or bodily expressive movements and verbal A/E for the prosodic utterance comprising the response cry and short assessments. These tend to produce brief communicative feedback effects while allowing the the cardholder to continue.

Ancillary questions, parallel then subjunctive assessments- This type of feedback would feature verbal recipient feedback extends beyond CP and would fall under C or A/E feedback functions. Potentially some co-speech gesture can occur with these assessments and questions counting as nonverbal C and A/E feedback. In response to these a cardholder would be expected to produce appropriate feedback back to the recipient, such as verbal and nonverbal CP (backchannels) and specifically verbal responses when answering ancillary questions that could be accompanied by co-speech gesture. This study aims to highlight the pattern of nonverbal feedback in which these occur for both recipient and cardholder.

Mid-sequence nods- It is expected there would a higher incidence of nonverbal CP than verbal CP in interactions featuring descriptions of bodily experience, a pattern different from other types of dialogue.

Second stories and formulations- Again, this type of feedback would feature verbal recipient feedback falling under the C or A/E feedback functions. These can be accompanied by co-speech gesture counting as nonverbal C or A/E feedback. Formulations may feature

co-speech gesture paraphrasing any original gesture produced by the cardholder; whether or not this can be regarded as simulative mimicry as discussed throughout is later covered in Section 8.1.1. Again, a cardholder would be expected to produce appropriate feedback to a recipient's contributions; including verbal and nonverbal C/P (backchannels) as well as affirmations, corrections to formulations or assessments of second stories- verbal A/E.

An additional overall hypothesis expects recipients should respond with stronger empathetic understanding by engaging in increasing levels of behavioural mimicry situations obligating empathetic communion- such as descriptions of negative experiences promoting altruism and sympathy (Preston and de Waal, 2002).

During clarification sequences cardholder and recipients are presented with issues in understanding, here it is expected participants will adapt their delivery to address these problems. Both parties are expected to take an active role at these points, recruiting more nonverbal signals overall. Cardholders will produce more gestures as they adapt their delivery when answering clarification questions. Recipients will collaborate on resolving the content of the problematic material addressed by the clarification question and respond with an increased use of descriptive gesture such as iconic and pantomime gestures. Gesturing to demonstrate what they do (or don't) understand of the content being described. This will see recipients making more contributions in this context than at in non-clarification sequences.

Due to the task structure it is expected the recipient would usually initiate a clarification sequence, asking for clarification to the description of bodily experience they are audience to. Although this will not always be the case as cardholders could indeed ask clarification question of recipient feedback such as second stories or ancillary questions.

7.3 Results

7.3.1 Embodied resources as empathetic feedback

Inter-rater reliability for the annotating feedback categories for both verbal and nonverbal feedback: contact/perception, comprehension and attitudinal or emotional was significant but comparatively weak.

Nonverbal feedback: Kappa = 0.201, p <0.001, n = 77, Standard Error = 0.04. Verbal feedback: Kappa = 0.411, p <0.001, n = 95, Standard Error = 0.04.

A closer analysis shows there is a weak agreement on comprehension feedback in both verbal and nonverbal feedback annotations. Excluding comprehension feedback, inter-rater reliability very good- with agreement on attitudinal-emotional almost perfect: Nonverbal: Kappa = 0.51, p <0.001, n = 64, Standard Error = 0.07. Verbal: Kappa = 0.94, p <0.001, n = 69, Standard Error = 0.04.

It appears annotators code contact/perception and attitudinal or emotional feedback much more consistently than comprehension feedback. As these feedback categories are incremental this may have occurred because coders found it difficult to determine, for example, when a nod signifies comprehension over contact/perception or when a 'oh yeah' is interpreted as attitudinal or emotional over comprehension, say if the coder detected an tone of surprise in the the vocalisation whereas a second coder could interpret the tone as indicative of demonstrating understanding. As these categories stack on top of each other incrementally, because comprehension falls between contact / perception and attitudinal or emotional feedback this type has more directions for error than the other two categories.

7.3.1.1 Nonverbal feedback frequency

Accounting for all annotated nonverbal signals, comprising of nonverbal feedback and descriptive gesture; Focussed comparison show recipients produce a large proportion of feedback signals (88%) whereas cardholders produce very few (12%).

The higher incidence of recipient feedback allows cardholders to judge how their description is being received. Feedback has been found to have an effect on narrative dialogues: the level of feedback produced by a recipient directly influences the fluency and effectiveness of the cardholder's description (see Section 3.2.8). The higher incidence of descriptive gesture for cardholders has been shown in Chapter 5- cardholders producing approximately five times as many descriptive gestures as recipients, and Chapter 6- people vary their posture more when speaking than as recipient. Again, this difference is expected as the cardholder's description of their recalled experience requires in-depth gesticulation. Therefore is it both expected for a recipient to produce a higher proportion of feedback signals (although see (Hadar et al., 1983))

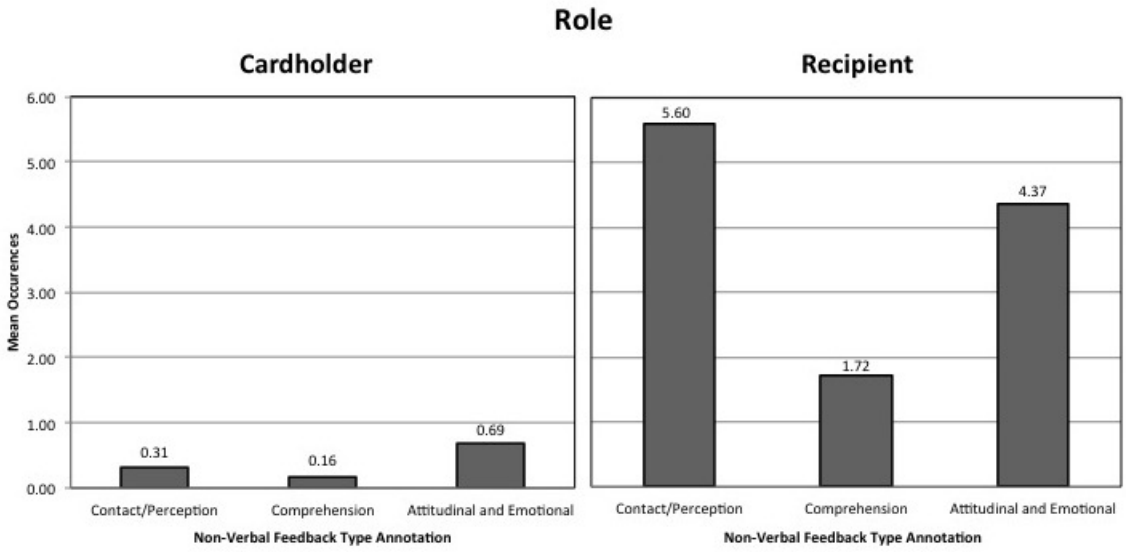
and conversely for a cardholder to produce a higher proportion of descriptive gesture.

Contact-Perception feedback comprise 42% of recipient's nonverbal signals but only 3% of cardholder's ($Chi^2_{(1)}$= 523, p = <0.001). Comprehension feedback comprise 13% of recipient's nonverbal signals and only 2% of cardholder's ($Chi^2_{(1)}$= 113, p = <0.001). Attitudinal or Emotional feedback comprise 33% of recipient's nonverbal signals and only 7% of cardholder's ($Chi^2_{(1)}$= 259, p = <0.001).

That being said, there will be a degree of overlap between descriptive gesture and nonverbal feedback; as the delivery of some nonverbal feedback signals will include descriptive gesture. These would mean the proportion of nonverbal feedback to descriptive gesture does not form part of a feedback signal will be somewhat higher for both cardholders and recipients. Descriptive gesture overlapping with feedback will not fall under the Contact-Perception category and descriptive gesture types do not incorporate the nodding backchannel expected under this feedback category.

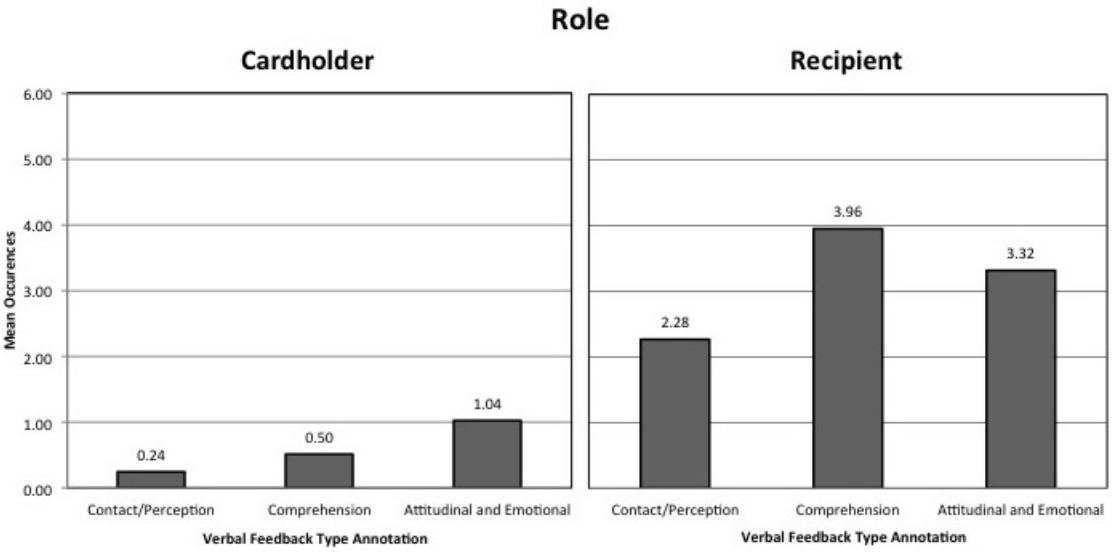
As well as absolute numbers the relative proportions of each nonverbal feedback type have different distributions ($Chi^2_{(2)}$= 43.5, p = <0.001). Cardholders make most use of Attitudinal and Emotional (60%) followed by Contact-Perception (26%) and Comprehension (14%). Recipients by contrast use Contact-Perception signals most (40%) followed by Attitudinal and Emotional (31%) and Comprehension (20%). Figure 7.1 shows the mean occurrence of nonverbal feedback signals for cardholder and recipient over each dialogue item, with recipients producing significantly higher levels overall. It could be said the Contact-Perception feedback signal, i.e. backchannels comprising of nodding, are prevalent in recipient feedback as they are produced so frequently along with a cardholder's speech- indicating each utterance has been understood so the cardholder knows to continue without interruption.

Figure 7.1. Incidence of nonverbal feedback signals



Contrastingly, Figure 7.2 shows a different distribution of verbal feedback signals. Recipients produce fewer Contact-Perception feedback signals than Comprehension and Attitudinal/Emotional verbal feedback signals. The higher incidence of Attitudinal/Emotional feedback suggests recipients are more engaged empathetically verbally than they are nonverbally. However, it is problematic to comment on the higher incidence of Comprehension feedback signals due to difficulties in coding this feedback type. Cardholders also produce more Attitudinal/Emotional verbal to nonverbal feedback suggesting a reciprocal verbal empathetic response to the recipient's verbal empathetic responses.

Figure 7.2. Incidence of verbal feedback signals



Chi^2 comparisons of feedback signals for Positive and Negative experiences by Recipient and Cardholder show no difference ($Chi^2_{(1)}$= 1.11, p = 0.29). A comparison of nonverbal Attitudinal and Emotional feedback also shows no difference between Recipient and Cardholder ($Chi^2_{(1)}$= 0.02, p = 0.90). Recipients do not appear to produce proportionally more empathetic feedback expressions for positive to negative experiences than a Cardholder.

7.3.1.2 Nonverbal feedback duration

A GLMM analysis of the average durations of each feedback behaviour with Role (Cardholder vs Recipient) and Valence (Positive vs. Negative) as a Fixed Factors and Participant as a Random Factor shows no reliable effect of Role ($F_{(1,339)}$ = 0.02, p = 0.88) or Valence ($F_{(1,339)}$ = 1.92, p = 0.17) and no Role × Valence interaction ($F_{(1,339)}$ = 2.6, p = 0.108). As previously seen in Chapter 5: valence of experience did not impact on the type of descriptive gesture used for either cardholder or recipient but did lead to recipients producing longer gesture when responding to descriptions of negative experiences. Unlike descriptive gestures, valence of experience has no effect on the distribution or duration of feedback signals.

7.3.1.3 Comparison of feedback patterns with control dyads

A full analysis of the control dyads in corpus 2 that is equivalent with the above analysis of the bodily experiences in corpus 1 is not detailed here. This is because firstly; there not a compatible positive / negative valence contrast in the control dyads, as the interactions are about informal current affairs of a relatively neutral content, they are not easily separated into positive or negative experiences by virtue of their topic. Secondly; the structure of the interaction has no role equivalent as there is no cardholder or recipient, so there is no task related reason that participants would not contribute equally. For this reason an analysis of feedback signal frequency, duration and distribution between interlocutors would not be directly comparable between the corpora. However, a useful comparison between corpora would be to analyse the distribution of feedback signals to give an insight into the distinguishing characteristics of how feedback differs for conversations about bodily experiences in corpus 1 and more typical conversational topics in corpus 2.

The control dyads rely more on verbal feedback types (65% of coded feedback) than the bodily experience dyads (47% of coded feedback) ($Chi^2_{(1)}$= 106.9, p = <0.001). Interlocutors produce more nonverbal feedback when talking about bodily experiences, this could be because recipients find it constructive to produce more embodied feedback, such as affect bursts (Schröder, 2003) or mid-sequence nodding to indicate an affirmation of cardholder's stance (Stivers, 2008).

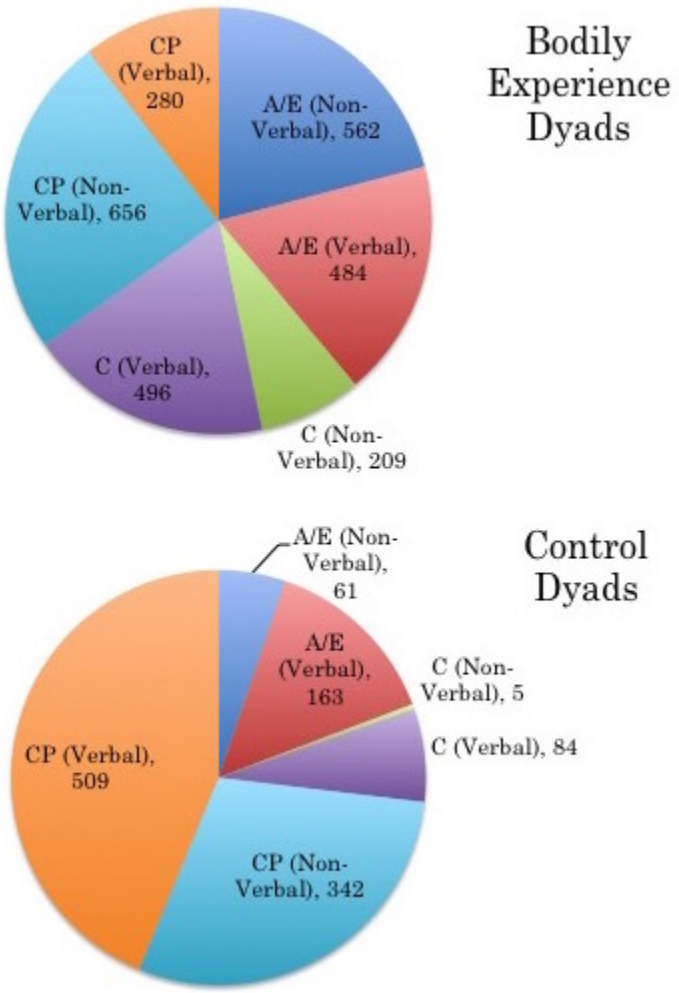


Figure 7.3. Distribution of feedback types between bodily experience dyads and control dyads

As the Pie Charts in Figure 7.3 show, while the control dyads predominantly provide Contact-Perception feedback, the bodily experience dyads make much greater use of Attitudinal/Emotional feedback and Comprehension feedback ($Chi^2_{(2)}$= 188, p = <0.001). The same pattern is also evident in the verbal feedback.

7.3.2 Using embodied resources in clarification dialogues

Inter-rater reliability for the annotating clarification questions and responses as perfect with Kappa = 1, p <0.001, n = 22, Standard Error = 0.00.

In analysis on all of Corpus 1, the bodily experience dyads, there are 4425 spoken turns tran-

scribed and 1349 descriptive gestures annotated. The majority of these (1205 or 89%) are produced by ardholders with only (144 or 11%) produced by Recipients. 96 of the total occur in clarification sequences. Clarification sequences are coded into two stages: The clarification question (CQ) and the response to the clarification question (R). Two separate dialogue roles are defined for clarification sequences: 'Speaker' for the participant currently talking and a 'Non-speaking addressee' for the participant not currently speaking, this is independent of the Cardholder/Recipient definition.

7.3.2.1 Gesture type and feedback

The combined results for both descriptive gestures and more generic forms of nonverbal feedback across all turns show interlocutors are active both when speaking and not speaking. An overall comparison of the frequency of nonverbal signals overlapping with the construction of a turn shows no reliable difference due to dialogue roles ($\chi^2_{(2)}$= 0.78, p = 0.38).

Table 7.1. Overall distribution of feedback and descriptive gestures by dialogue role

Non-speech signals:	Descriptive gesture	Feedback
Speaker	2113 (79%)	550 (21%)
Non-speaking addressee	553 (21%)	1975 (78%)

As Table 7.1 shows, although Non-speaking addressees are active they do not contribute to the same types of nonverbal signals as Speakers. Speakers produce around 80% of descriptive gestures overlapping with their own turn whereas Non-speaking addressees produce around 80% of the nonverbal feedback overlapping with the Speaker's turn. This fits the general pattern reported previously (see Section 7.3.1.1) and in the literature attentive Non-speaking addressees provide frequent 'backchannel' feedback to signal continued attention and interest and to support the Speaker in the construction of a turn (see Section 3.2.8). These results also suggest Speakers provide some nonverbal feedback signals to the Non-speaking addressee while they are talking.

Table 7.2. Distribution of feedback and descriptive gestures by dialogue role in clarification sequences

Non-speech signals:	Descriptive gesture	Feedback
Speaker	128 (69%)	57 (31%)
Non-speaking addressee	94 (49%)	99 (51%)

In the corpus a total of 194 turns were coded as clarification questions and 183 as responses; this is in line with Purver et al. (2003) analysis of clarfication questions accounting for 4% of dialogue turns in ordinary dialogue. Of these the majority of clarification questions are initiated by the Recipient (166 = 85%) and the majority of responses are provided by the Cardholder describing the experience (87%).

The prediction descriptive gestures will be more frequent overall during clarification dialogues is not supported overall, only slightly more clarification sequence turns overlap with descriptive gestures (26% vs 20%). However, the pattern of gesture frequency between dialogue role changes; Speakers gesture less frequently during clarification sequences than otherwise (dropping from 33% of turns overlapping with gestures to 26%) whereas Non-speaking addressees gesture more than double their usual descriptive gesture frequency during clarification sequences rising from 7% to 19%.

As Table 7.2 shows, the general pattern for Speakers and Non-speaking addressees shown in Table 7.1 changes substantially during clarification sequences. Non-speaking addressees become substantially more likely to produce descriptive gestures in overlap with the Speaker's turn (increasing from 21% to 49% of their gestures $\chi^2_{(1)}= 0.734$, p <0.001) and Speakers themselves become more likely to produce nonverbal feedback in overlap with their own turn (increasing from 21% to 31% of turns, $\chi^2_{(1)}= 10.6$, p <0.001). This supports the prediction non-speaking addressees make more use of descriptive gestures during clarification sequences. As Speakers have 'epistemic access' to the source of what is being clarified their feedback could be argued to be based around confirming or acknowledging appropriate recipient feedback demonstrating understanding.

7.3.2.2 Hand movements during clarification questions

To assess whether speakers normally move their hands more than non-speaking addressees in this corpus, a Generalised Linear Mixed Model (GLMM) analysis of the average frequency of hand significant movements (as defined in Section 4.5.2) for each participant during all dialogue turns compared the hand movements with Speaking as a Fixed Factor (Yes vs. No) and Dyad as a Random Factor. This shows a main effect of Speaking (Yes or No) ($F_{(1,518)}$) = 96.1, p <0.001). Speakers move their hands approximately 25% faster (5.42) than Non-speaking addressees (4.03) across all turns.

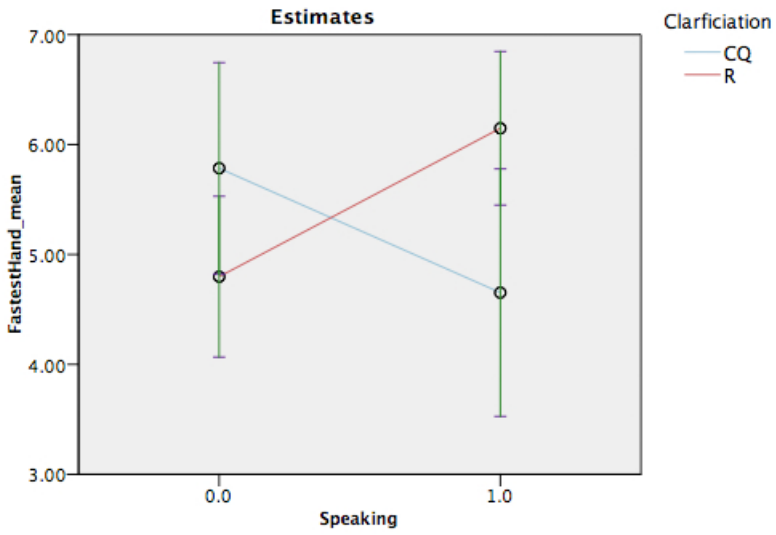
A second focused GLMM analysis of hand movements during clarification dialogues for each trial with Dyad as a Random Factor and Speaking (Cardholder or Recipient), Clarification Stage (Clarification Question or Response) and Clarification Stage × Speaking interaction shows no simple main effects of Speaking ($F_{(1,143)}$) = 0.12, p = 0.73) or Clarification ($F_{(1,143)}$) = 0.58, p = 0.45) but a significant Clarification Stage × Speaking interaction ($F_{(1,143)}$) = 14.24, p <0.001). The interaction is illustrated in Figure 7.4.

In this figure the red line and the dots it connects refer to the clarification question (CQ), the blue line and associated dots refer to the response (R). The FastestHand mean axis is the speed of the fastest hand in seconds. The Speaking axis labels 0 and 1 as Cardholder and Recipient.

7.4 Discussion

Overall, the hypothesis the organisation of feedback to dialogues featuring descriptions of bodily experience would take a different pattern to other types of dialogue is observed in these results. Here, feedback strategies for the control dyads and bodily experience dyads differ. The control dyads, featuring discussions of current affairs tend to produce a large proportion of both verbal and nonverbal backchannels, a small proportion of verbal attitudinal and / or emotional feedback and hardly any other type. This could perhaps suggest for these interactions interlocutors offer ongoing feedback their partner is being heard as well as contributing their opinion via a verbal agreement or disagreement. Contrastingly, the bodily experience dyads produce a different pattern. They produce a smaller proportion of nonverbal backchannels and an even

Figure 7.4. Estimated marginal means of the speed of the fastest hand (in seconds) for the interaction between clarification stage and speaking.



smaller proportion of verbal backchannels than the control dyads, with a higher proportion of both verbal and nonverbal attitudinal / emotional feedback. The higher use of both verbal and nonverbal attitudinal / emotional feedback in the bodily experience dyads demonstrate interlocutors are more empathetically engaged when they are talking about bodily experiences, likely due to the affective subject matter. As mentioned previously, Heritage (2011) asserts a description of a bodily experience the cardholder has sole epistemic access obligates the recipient to offer affirmation of the nature of the experience and affiliation with the cardholder's stance towards it. Therefore the attitudinal / emotional feedback observed in the bodily experience dyads could be said to fulfil this moral obligation and contribute towards creating moments of empathic communion.

The prediction interactions featuring accounts of personal experience will see recipients respond by displaying a high proportion of nonverbal empathetic feedback such as response cries or affect bursts than other types of dialogue is compatible with the analysis as recipients produce a high incidence of C and A/E nonverbal feedback would encompass this type of feedback. In addition, the prediction recipients respond with empathetic verbal feedback such as ancillary questions, parallel and subjunctive assessments, second stories and formulations is also compatible as the high incidence of verbal C and A/E feedback in contrast to the control dyads. Co-speech gesture accompanying these verbal responses will also contribute to boost the proportion of nonverbal C and A/E feedback seen in these results. The prediction recipients would respond

to a cardholder's description with mid-sequence nods as opposed to verbal backchannels to show affiliation with a cardholder's stance is supported by these results by the high incidence of nonverbal C/P feedback from the recipient, showing over twice the amount of occurrences than verbal C/P feedback. This type of empathetic feedback appears to be the type relied on the most. It could be said this is an influence of role, the description of the experience must take precedence, verbose and substantive forms of feedback being offered only when the description has been completed by the cardholder, resulting in a lower frequency.

Despite the cardholder's description taking precedence over the dialogue the results show cardholders do produce a degree of feedback reciprocally to the recipient's feedback contributions. The presence of verbal and nonverbal C/P in the results is compatible with the hypothesis cardholders produce verbal and nonverbal backchannels back to the recipient's more verbose and substantive feedback responses. The effect of role means recipients contribute less dialogue than cardholders and so cardholders have less to feedback to, explaining why cardholders use less backchannels. The results shows cardholders' rely on verbal and nonverbal A/E more than backchannels. This is compatible with the suggestion cardholders answer ancillary questions by demonstrating a stance or to communicate affect, implying cardholders do respond with affirmations, corrections to formulations and deliver assessments to second stories - all resulting in occurrences of verbal and nonverbal C and A/E feedback. That a cardholder relies more on this type of feedback rather than backchannels could be indicative of a sense of ownership over the account and the delivery of the talk- cardholders tending to offer their stance and point of view rather than producing continuer like feedback encouraging a recipient to hold the floor.

The hypothesis recipients would respond more strongly to accounts of negative experiences due to an expected tendency towards altruism and sympathy (Preston and de Waal, 2002) was not supported by this analysis. The prediction interlocutors would produce more nonverbal empathetic responses, such as affect bursts, to communicate understanding and mutual recognition of the cardholder's experience is not observed as there is no influence of valence on the pattern of feedback for either cardholder or recipient. The pattern of empathetic responses appear to follow the same interactional strategy whether the topic is positive or negative. This could be because the recipient doesn't judge the situation sufficiently serious for sympathy of negative experiences to be warranted as the incident being described was now resolved.

Turning to the analyses on clarification sequences, at junctures exhibiting problems in understanding such as this it is expected interlocutors will use a distinctive pattern of nonverbal resources where both speaker and non-speaking addressees collaboratively adjust their contribution in attempt to move forward with the conversation.

Overall these results shows during the clarification dialogue as a whole the speed of speaker's and non-speaking addressee's hand movements are not reliably different. However, they vary across the different stages of the clarification sequence. When the clarification question is asked by a speaker the non-speaking addressees move their hands more than speakers. During the response speakers answering the question move their hands more than non-speaking addressees audience to the answer. In other words, when a clarification is posed non-speaking addressees are moving their hands more than speakers but not when it is answered. Non-speaking addressees more than double their use of descriptive gesture. Although speakers still produce the most descriptive gesture they also produce more feedback. In line with predicted, both collaboratively recruit additional, multi-modal resources to overcome problems for mutual understanding.

As noted above, clarification questions in this task are mostly, although not always, posed by the recipient which suggests the cardholder describing the experience does not immediately suspend their hand movements when a question is posed about something they've just said. Normally speakers reduce their hand movements after their turn.Sikveland and Ogden (2012) describes post-stroke gesture holds held over clarification questions as 'pending completions', functioning as a resource for the speaker to display there remains an outstanding issue in shared understanding. During their response, again usually produced by the cardholder describing the experience, they continue to move their hands more quickly than the recipient. This suggests 'ownership' of the content being questioned has impact on the organisation of hand movements at these points (Heritage, 2011), highlighting the assymmetrical nature of epistemic access and how this is addressed in conversations about bodily experience.

Gestures produced by non-speaking addressees are interesting because most models of gesture production are based on the components of the cognitive system produce them being connected bodily and mentally (Gentilucci and Dalla Volta, 2007). Facilitating word searches (Krauss, 2000) and conceptual planning (Alibali, 2010). Here, the collaborative use of gesture and speech concurrent between interlocutors does not fit with these models.

7.5 Summary

This chapter analysed the organisation of feedback signals for interactions featuring descriptions of bodily experience, predicting interlocutors will respond and engage with these accounts in a strategically unique manner than to other types of dialogues. Following Heritage (2011) and (Kupetz, 2014), the pattern of feedback in these results demonstrates recipients produce the majority of feedback signals whose function is appropriately empathetic to a cardholder's talk, reciprocally cardholders produce feedback appropriate to their role. Feedback strategies between interactions of accounts of bodily experience differed significantly from interactions on current affairs. Recipients adjust their feedback strategy and use their embodiment in ways appropriate to the material, without replicating or recapitulating it. The hypothesis interlocutors would produce a stronger empathetic response to accounts of negative experiences was not supported in this analysis, interlocutors follow the same interactional strategy regardless of valence.

During clarification sequences speakers and non-speaking addressee nonverbal behaviours tend to merge. Non-speaking addressees producing more than twice as many descriptive gesture in overlap with a speaker's turn than otherwise. This underlines the collaborative nature of conversation, the strategic importance of nonverbal resources for sustaining mutual understanding and the critical role of clarification and repair in working towards intersubjectivity.

In conclusion, this analysis echoes the results from chapter five; that descriptive gesture are governed by higher cognitive processes that are interactionally involved and strategic communicating a message or demonstrating understanding. and chapter six; that posture appears to be autonomous, or directed by constraints related to the act of communicating. In this analysis interlocutors are seen to adjust their nonverbal feedback strategies appropriately to the content, without imitating it. Nonverbal contributions are primarily constructed to aid the collaborative, sequential and intersubjective progression of dialogue.

Chapter 8

Recipient performance as an interactional strategy for intersubjectivity

The previous analyses show interlocutors use their embodied resources differently in different interactional circumstances: Corpus 1, designed to provoke affect-laden descriptions of bodily experience that elicit empathetic responses features a pattern of descriptive gestures that focus on either inward sensation (abstract descriptive gesture), expressive reactions to events/experiences (pantomime gesture) or provides a situational description (iconic and deictic gesture) to contextualise these experiences or events. Corpus 2, designed to emulate informal conversations about topics of interest to each party but generally unrelated to a personal experience feature fewer gestures that convey affective and emotional content. The difference in gesture type distribution in these corpora is indicative that the corpus based around bodily experiences is successful in provoking empathetic interactions as both parties use their embodied resources more specifically to communicate about affective experiences. As this corpus revolves around one participant (the cardholder) describing a recalled experience to their interactional partner (the recipient) the situation is akin to a storytelling interaction and this is reflected in the higher frequency and longer duration of descriptive gestures made by the cardholder.

Recipients still employ their embodied resources by producing feedback expressions and the relative pattern of gesture types used by cardholder and recipient is not significantly different.

Although the global distribution of gesture types does not differ there is still the possibility of a behavioural mimicry effect to be present in this particular set of gesture types. The perception-behaviour link mechanism, that people automatically behave as they perceive due to thinking and doing sharing the same neurological representations, has been deemed conducive to prosocial effects between interlocutors (see Section 3.2.4). As these studies have not measured mimicry of descriptive gestures that contribute to the message being conveyed in conversation (rather than self-adaptor such as foot-tapping or face touching), evidence of manifest mimicry behaviour of affective expression would strengthen the argument that the direct link between perception and behaviour prompts recipients to simulate the specific affective or emotional state of the cardholder. However, such a global measure only suggests that recipient's and cardholder's gesture type distribution is different. These gestures could nonetheless be of same content and therefore counted as behavioural mimicry, similar in morphological form and meaning.

Posture congruence results don't show the presence of behavioural mimicry in this corpus. Although a similar pattern of gesture types are produced by cardholder and recipient, their overall postural similarity is not above chance levels, even over a range of possible temporal offsets. A problem with these results is that they average similarity between multiple body parts, so the possibility remains that a mimicked expression performed by one body part similarity is cancelled out by another body part that is dissimilar, or the mimicked hand gesture performed by the opposite hand (mirrored). Size and velocity of a gesture could differ while still retaining similar morphologically form and meaning which would count as mimicry- whilst not being picked up by the postural congruence measure used here.

Bavelas et al. (2000); Bavelas (2011) show that recipients are active in collaboratively co-narrating by producing appropriate feedback that helps the narrator progress their story. This is evident in the difference in feedback type distribution between this corpus of description of affect-laden experiences and the more symmetrical corpus featuring conversations of topical interest. The higher proportion of non-verbal attitudinal or affective feedback in the affect-descriptions shows a marked increase in empathetic responses such as affect-bursts and motor mimicry, a further indication there is more inclination to respond empathetically.

Thus, while there appears to be no direct evidence of automatic copying of body movement in this corpus there is clear evidence of active engagement. This chapter asks what the inter-

actional contexts in which recipients employ their nonverbal resources to either engage with or display empathetic understanding in response to affect-laden tellings of bodily experiences. Specifically, the analysis investigates if and how recipient's nonverbal contributions are cooperatively organised sequentially alongside the cardholder's descriptive contributions in order to collaboratively advance to an intersubjective understanding between them.

This chapter takes a closer moment-by moment analysis of six sequences that include nonverbal recipient feedback that appear to be examples of empathetic engagement. The examples were chosen by scanning each incidence of recipient nonverbal feedback and selecting recipient expressions that were closest probable instances; firstly, of behavioural mimicry and secondly, congruent behaviour in the perspective of the cardholder. Behavioural mimicry is defined in the selected sequences as a bodily expressions that are similar in morpholigocal form and represent a similar meaning. Generally, the imitated expression and the imitation occur instantaneously or withing thirty seconds of one another (see figures attached to examples). Congruent behaviour in the perspective of the cardholder is defined as what Bavelas et al. (1986, 1987, 1988) term as motor mimicry (see Section 3.2.7.2)- the performance of an expected expressive behaviour in the perspective of another.

Each example was transcribed in ELAN using guidelines from Atkinson and Heritage (1985) transcript notation for conversation analysis, each utterance analysed in context with the expressions that accompany them in sequence, for example nodding is transcribed were it occurs temporally in the excerpt. The analysis was conducted in ELAN with a video, audio and the transcription annotation synchronised on a timeline. The sketches and transcripts are for the reader only and were not used in the analysis. The sketches are produced by tracing over still images from the video so are, if reduced, representations rather than interpretations. Each example features a sequential description of the excerpt, with some cursory reference to general interactional events that occur, then an in-depth discussion of the contribution of embodied resources such as gestures, postures and expressions to empathy and intersubjectivity.

8.1 Adapting expressions to demonstrate intersubjectivity

Pantomime gestures provide a rich emotive context to a description of an event or experience (see Section 4.6.5.3). By performing or acting out their reactions toward an event (Or in the

Table 8.1. Example gesture types compared between cardholder and recipient

Example	Cardholder	Recipient
1	Pantomime	Pantomime
2	Pantomime and Iconic	Pantomime
3	Pantomime and Deictic	Pantomime
4	Iconic, Deictic and Pantomime	Pantomime
5	Pantomime	Pantomime
6	Pantomime and Deictic	Pantomime

case the story is second hand, their characters reaction) a speaker gives their recipient the ability to judge the experience of the Other through direct observation of their behaviour. Pantomime gestures give supplementary information about the embodied aspect of the situation:- not only through *situational cues* about an experience, such as which foot, what toe and what movement led the person to stub their toe; but also *emotional / affective cues* gives information about the intensity of the pain, or an emotional response such as anger or embarrassment. It can be argued that by employing non-verbal resources in this way leaves less room for mis-interpretation than verbal resources in descriptions of embodied experiences (such as pain). Pantomime gestures take advantage of embodied resources to give a level of detail that isn't as readily available in words, if at all. Even the most colourful of language cannot give as much specificity as a particular quality of movement, such as a grimace, so directly. Pantomime gestures tend to be exaggerated versions of behaviour, this is in line with the argument that these expressions are capitalised on to serve a communicative function (Bavelas, 2011).

Husserl's concept of 'pairing' suggests that by being similarly embodied we have the capacity to understand each other's experience through our manifest behaviours. Using pantomime gesture in a description of an event of experience gives the recipient access to this behaviour in a re-telling setting. Through 'pairing' an understanding of the experience behind the behaviour is available without the need to analogise or inwardly relive each other's experience through simulation:- I do not need to simulate your breathing in order to understand that, like me, you also breathe, I know this because we are similarly embodied. Note that this concept of 'pairing' doesn't assume that we are identical, just similarly embodied. Therefore 'pairing' does

not provide an infallible understanding of an expression- only that it is possible to make the interpretation that the experience behind the pantomime gesture is of a certain type and in this way gives access to the speakers' emotional or attitudinal stance. An additional consideration is that this interpretation is also governed by other influences, such as the social factors like discussed by Schutz (see Section 2.4).

The question of concern here is how do people actually use pantomime gestures in interaction? How are they deployed as part of a specific sequence of turns? As pantomime gestures in re-tellings of experiences perform the original behaviour in these stories- do recipients imitate them as a simulative mechanism to understand the original experience? Recipient responses that embody the perspective of a speaker by imitating their behaviour or would-be behaviour can be seen in the phenomena of motor mimicry (see Section 3.2.7.2), contrastingly Bavelas and colleagues make the case that this behavioural event (i.e. wincing at the pain of another) is not a by-product of an intrapersonal processes such as a simulative mechanism but an attempt to *display* similarity to the speaker. Here motor mimicry has the function of communicating 'I am like you' (Bavelas et al., 1986, 1987, 1988).

The following examples explore the hypothesis that imitation is not indicative of a simulative mechanism to understand the experience being described but functions to strategically attempt to display understanding, sympathy or affiliation.

8.1.1 Example 1: Pantomime adaption

In the following example the cardholder (In all figures in for example 1 the cardholder is positioned on the right of the sketch) is describing why she enjoys yawning. In Figure 8.1 she sets up the situation with a statement 'I like yawns' in line 1, following this she hesitates while continuing her justification, the sketch shows her raising her arms to emulate the preparation motion of stretching out as if in the act of yawning but stops short in line 2 'Er er', as if considering how to articulate the aspect of yawning she likes. Filled pauses containing utterance like 'Er' 'Uh' are common signals of self-repair, indicating that a cardholder is having trouble formulating or articulating their message. Healey et al. (2013) find that recipients respond more strongly by producing more backchannels to turns than contain self-repairs (Healey et al., 2013). This is reflected in this excerpt in line 2 where the recipient recognises the trouble

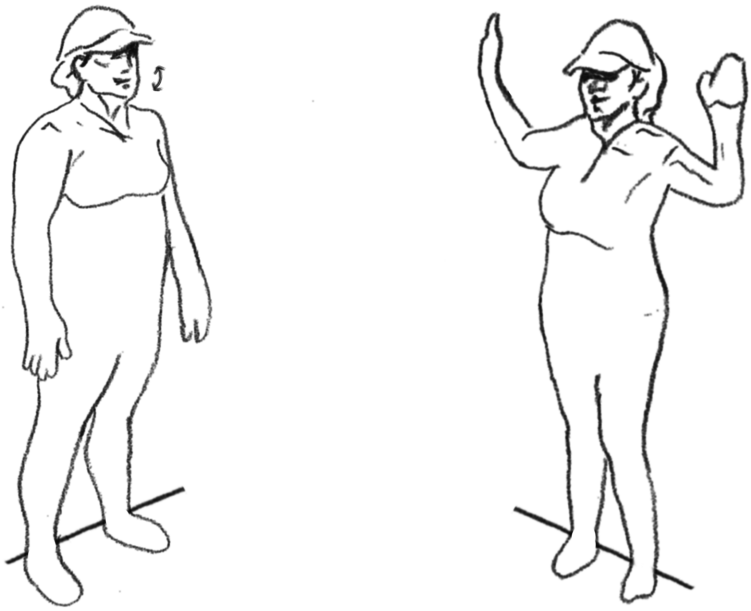
the cardholder is having and recipient nods to demonstrate she has received the message, the nodding is indicative of a marker signifying for the cardholder to continue (Cerrato, 2002). Additionally, as discussed in Section 7.1 Stivers (2008) shows a mid-sequence nod positioned where the cardholder offers a telling element that provides insight into their stance, i.e. the 'Er I like yawns', displays affiliation with the cardholder's position.

In Figure 8.2 the cardholder moves back into the preparation motion of fully stretching out and says with a false start at line 3 'I like- I like the stretchy sensation'. It is clear that the cardholder knows the embodied actions of yawning that she likes as she moves into the action as she says 'I like yawning' at the beginning, but falters in articulating the 'stretchy sensation' of yawning that she likes. The recipient responds with 'Yeah yeah' in line 4. The utterance is accompanied with a pantomime gesture of her own- recounting the action of stretching back to the cardholder. It is clear she is not simply imitating the expression of the cardholder because she produces her own version of the action. The cardholder raises her arms to prepare for stretching out from the elbow, whereas the recipient stretches her hands out from her wrists. Bodily, the enactment produced by the cardholder (stretching out from the elbows) is different from the enactment produced by her recipient (stretching out from the wrists) but conceptually both movements are actions derived from the same type of experience. It wouldn't be possible for her to produce her own version of the enactment if she didn't already possess a sense of the type of yawning experience being described. If a simulation of yawning was required for her to understand she would not know how to modify the expression in her own way. For this reason, she has not imitated her interactional partner as a simulative mechanism to understand the act of stretching.

Instead, the adapted stretch serves a strategic communicative function:- the knowledge that the recipient has an understanding of the experience the cardholder is trying to communicate is demonstrated by her performing what could be referred to as a *paraphrase* of her stretch appropriate to the type of experience being described for the cardholder, serving as proof she is successfully following the account. It is worth noting that if it was the case that the recipient did in fact produce an imitation of the action, the act of copying perceived activities alone provides no evidence the recipient has advanced further than the mere perception of a movement. Whilst this can certainly demonstrate her ability to imitate the movements she sees, it cannot demonstrate an understanding of the type of experience associated with the

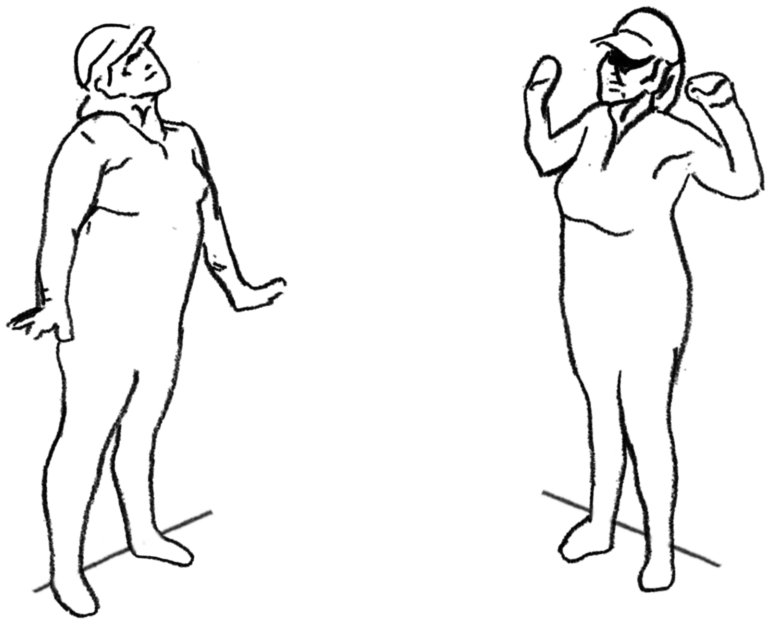
action, whereas her own variation of the movement serves as a more powerful shadowing tactic demonstrating an understanding that is specific to a particular experience type.

Figure 8.1. Stretching sketch



1. CH: Er I like yawns
   (.)
2. CH: [Er er]
   NCH: [(nods)]
   (.)

Figure 8.2. Stretchy sensation sketch

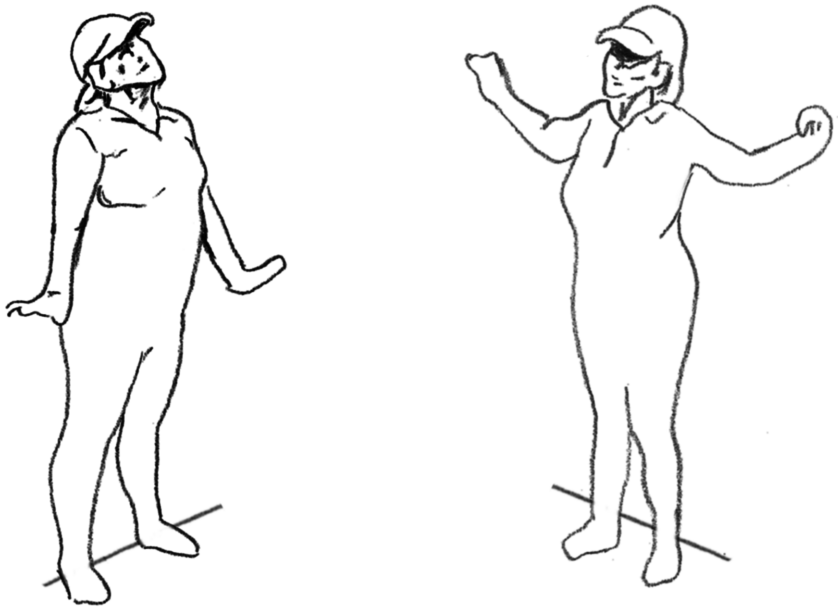


3. CH: I like- I like the stretchy sensation

(0.2)

4. NCH: Yeah ye::ah=

Figure 8.3. Stretching together sketch



5. CH: =the one in the morning
(0.3)
6. CH: er
(.)
7. CH: usually
(0.3)
8. CH: I like those

This stretch functions as feedback for the cardholder designed to *demonstrate* understanding without interrupting the cardholder's description instead of plainly stating it. As the recipient performs the gesture in a toned down manner (her arms hardly leave her sides), the gesture is far from a 'bid for the floor' such as the turn grabbing / self-selecting pointing gestures found by Mondada (2007), or the palms up turn grabbing gestures found in Streeck and Hartge (1992), but similar in form to the style of assessments found in Goodwin (1986). These are defined as verbal responses like 'Oh wow' or simple sounds like 'A::h' from recipients that assess the content of the on-going talk, sequentially displaying understanding without interrupting the turn in progress. In this way they are close in function to more standard continuers (Cerrato, 2002) but provide additional feedback.

The demonstration of understanding is shown on the periphery; giving the cardholder space to move into the full stretch in Figure 8.3 at line 5 where the two gestures overlap. Now she knows she can proceed the cardholder gives additional contextual details - 'the one in the morning'. The recipient's variation of the stretch is terminated when the cardholder completes her stretching out, and this completion is marked by the recipient with a nod.

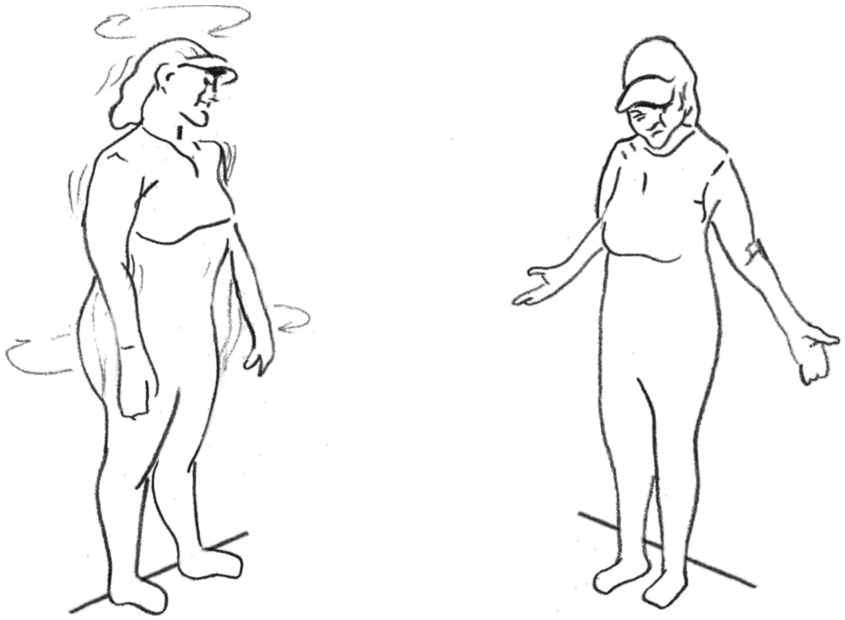
Healey et al. (2013) suggest that it is at points where conversation threatens to go off-course that are the most significant for coordination, the presence of self-repair in this excerpt could lead to the recipient feeling more compelled to utilise feedback behaviours from backchannels to expressions that demonstrate her understanding in order to help the cardholder along.

The communicative function of these types of adapted expression is further shown when the conversation turns back to the 'morning yawn' in the following excerpt 56 seconds later in the conversation. The cardholder's utterance in Figure 8.4 at line 1 'Er yeah but the one in the bed I think it's the best one because' is closely followed an agreement from the recipient in the transition space with 'Yeah ye::ah' in line 2, the second 'yeah' pronounced with an elongated 'air' vowel sound with a falling intonation. The utterance is not only verbally affirmative, but distinctively imitative of the sort of sigh produced when yawning. The sketch shows the recipient also performs a pantomime gesture overlapping with the second elongated 'Yeah' which could be described as emphasising the bodily reaction to the stretchy yawn, a wiggle which overlaps the cardholder's utterance in line 3. The gesture is unlike a typical turn grabbing gesture as again her hands remain by her sides and occurs outside of the transition space but overlaps the cardholder's utterance in line 4 (Streeck and Hartge, 1992; Mondada, 2007). In addition, the

movement doesn't stop short of the stroke phase of the gesture to wait for a turn but continues over the top of the cardholder's utterance so cannot be in preparation for articulation (Harrigan, 1985). An enactment of the body relaxing as a result of the stretch of the yawn that goes beyond an adaption and describes a new aspect of the yawning experience. What makes this contribution unlike behavioural mimicry is that it conveys an aspect of the experience that is not initiated verbally or gesturally by the cardholder. It's a different interpretative description as the recipient's expression reworks the original display firstly by the addition of a vocalisation, but she also performs her own style of representing the experience. It is not an imitation of an action performed by the cardholder but constitutes an *extension* or an *elaboration* of the cardholder's talk. It is through this elaboration that the 'wiggle' representation provides evidence of a broader understanding of the experience being described. Again the 'wiggle' doesn't disrupt the cardholder's talk but provides a back-channel style contribution rich in information associated with the experience being described.

That the recipient doesn't imitate the cardholder's descriptive gesture could be down a respect for the cardholder's 'ownership' of the experience being described. Heritage (2011) observes that such types of bodily sensation are only directly available to whomever experiences them and for this reason when cardholders describe this experience their recipients must deem them as having superior epistemic access. When responding in an interaction like this recipients must balance respecting a privileged access to the experience with the moral obligation to show affiliation and *empathetic communion* (Heritage, 2011; Kupetz, 2014). By adapting the expression the recipient avoids claiming direct access to the specific experience being recalled by the cardholder but demonstrates understanding and affiliation with the *type* of experience.

Figure 8.4. Wiggle sketch



1. CH: Er yeah but the one in the bed I think it's the best one because
   (0.2)
2. NCH: [Yeah ye::ah=]
   [(wiggles)]
3. CH: =I like the stretching

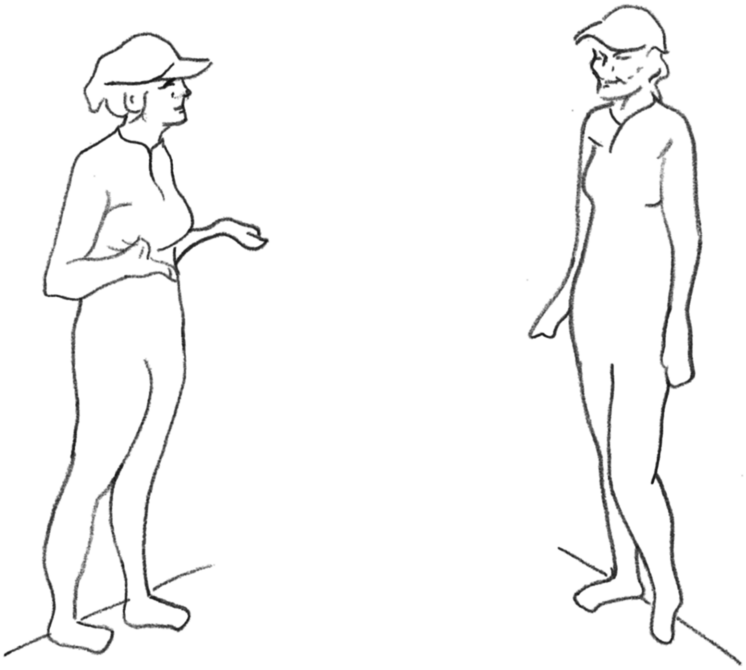
8.1.2 Example 2: Elaborating on descriptive expressions in the presence of disfluencies

In this excerpt the cardholder (In all figures in for example 2 the cardholder is positioned on the left of the sketch)has previously been talking about a recent backache that was caused by the repetition of a particular movement until her back muscles were tired and she uses pantomime gestures to enact the movements that caused the pain. The cardholder continues this previous talk in Figure 8.5 from line 1 'Just kind of one of those' after which there is a very small pause and she says 'pain that's there' in Figure 8.6 at line 2. The sketch shows this utterance is accompanied by a gesture where she places her left hand palm up to the left of her gesture space and circles her right hand on top as if rubbing a shape onto a surface, indicative of the size and shape of the area of pain.

Line 2 is followed by a long pause of 0.7 seconds before she says 'but doesn't-' in Figure 8.7 at line 3 adjusting her gesture to enact the sort of posture adopted in reaction to a backache, tensing and then rolling her shoulders enacting a pain response. Another longer pause of 1 second follows line 3. The cardholder completes her sentence with a false start in line 4 'it's just tir- a ti:red pain'. These particularly long unfilled pauses [1] and false starts are further instances of self-repair (see Section 8.1.1); indicating a problem in expressing the 'kind' of pain she sets up in the previous talk. Here, the recipient nods and producing a vocal continuer 'mm' during the long accentuation of the 'i' vowel sound in the word 'ti:red' showing further indication that recipients attempt to help to move the conversation on by responding more strongly to turns that contain disfluencies (Healey et al., 2013). At the same time she starts to raise her hands palm up facing the cardholder as if to prepare to talk, similar to what Streeck and Hartge (1992) describe as a self-selecting turn grabbing gesture. This is also in line with Harrigan (1985) finding that recipient body movements occur significantly more frequently during turn-requesting phases; suggesting the recipient intentionally uses gesture to signal her bid to take the floor (Harrigan, 1985).

[1] Average intra-turn pause length is approximately 0.5s (Fors, 2011).

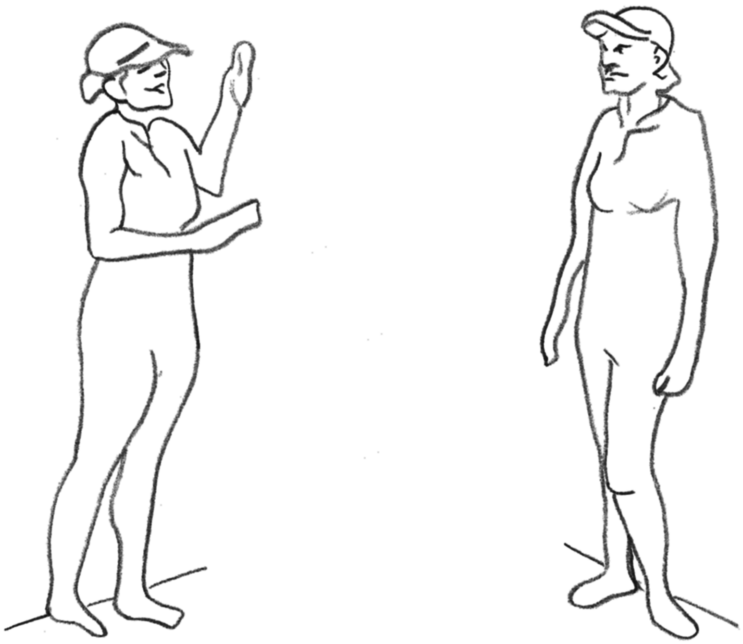
Figure 8.5. One of those?



1. CH: just kind of one of those

(.)

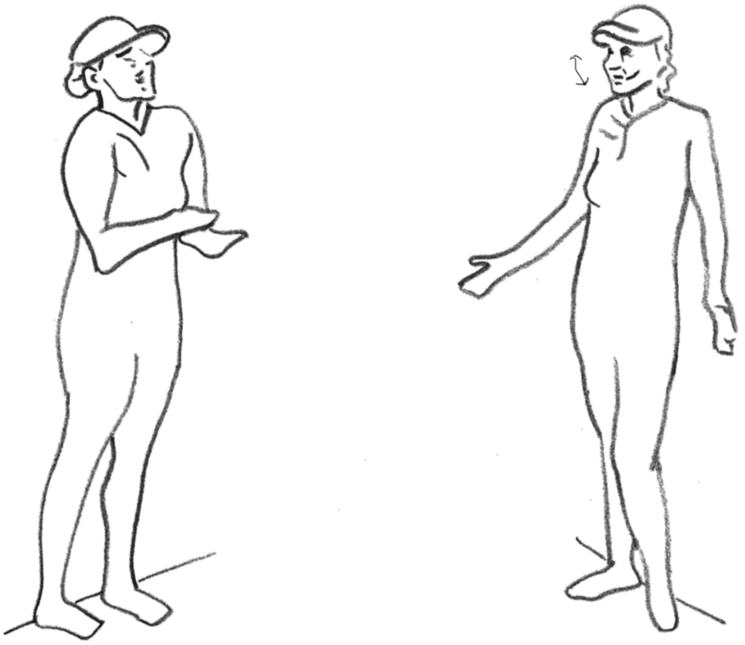
Figure 8.6. Pain movements



2. CH: pains thats there

(0.7)

Figure 8.7. Backpain



3. CH: but doesn't

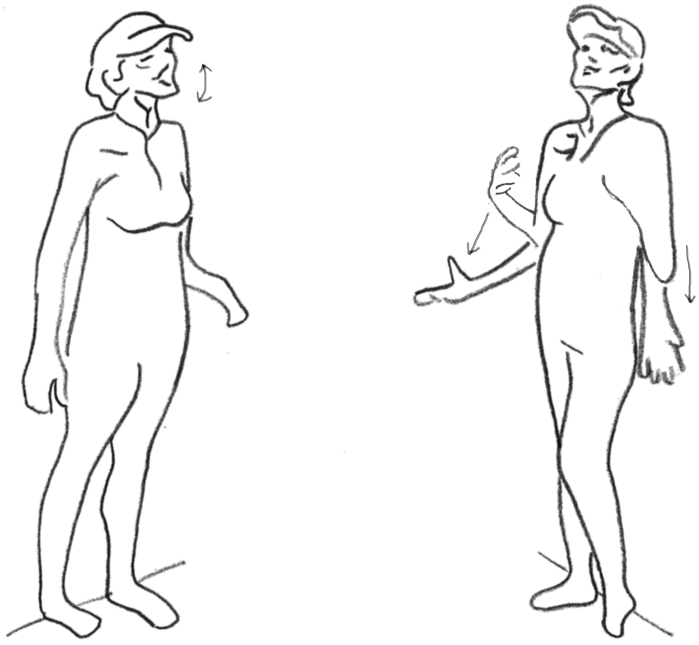
(1.0)

4. CH: it just tir- a [ti:red pain], but [not]

NCH: [mm mm]

NCH: [(nods)]

Figure 8.8. The muscles are like...



5. NCH:[Yeah] and the muscles are just like [a:::h]=
   CH: [(nods)]
6. CH: =Yeah

Duncan (1972) counts a stressed syllable or 'drawl' on the end of a clause as a turn-yielding cue so the recipient's turn-request as this point in the sequence is intentional and down to the cardholder producing an elongated 'ti::red pain' at the end of her clause. The cardholder starts to make an additional clause 'but not' but upon noticing the recipient's bid for the floor she ceases her turn while dropping her arms to her sides in order to hear the recipient's contribution at line 5 in Figure 8.8. The termination of her gesture another turn-yielding cue (Duncan, 1972)

The recipient says 'Yeah and the muscles are just like a:::h', the 'a:::h' a relaxing sigh. At this point an abstract descriptive gesture is produced:- the recipient raises her hands to chin height and then moderately slowly releases them out in front of her, tilting her chin up indicating the transition of a tense posture to a relaxed one. The gesture is not intended to be wholly complementary to the speech in the recipient's turn, but is an integral component, completing the expression the multimodal behaviour intends to articulate. Both vocalisation and bodily behaviour together contribute to the whole expression, each modality considered as concurrently disambiguating each other, rather than predominantly considering the use of gesture as a resource to disambiguate speech (Holler and Beattie, 2003).

Nonetheless, it is evident this gesture is describing the character of the experience of having tired muscles but its emphasis is different from both types of gestures that the cardholder produces. The recipient's contribution presents an alternative interpretation the character of the experience, without imitating or adapting the gestures she perceives. This description is different from the two examples above, as the recipient turns from acting out the experience to describing the experience itself:- what the muscles *feel* like. At this point it is necessary for her to extend her presentation with the clarification that she is talking about the muscles by saying 'and the muscles are just like', otherwise it could seem as though she were performing a reaction to the experience in the perspective of the cardholder like the other examples. This assessment does interject the cardholder's description and for this reason there is less resemblance to a continuer or back-channel type feedback in this instance.

Why the recipient felt compelled to interject the cardholder's description with her own assessment may be a response to the presence of long pauses in the middle of the cardholder's clauses in these previous turns. During which the cardholder attempts to describe the sensation using a couple of different types of gesture: pantomiming a bodily reaction to pain as well

as describing shape and size and hesitating between utterances. At this point the recipient interjects to assure the cardholder of her understanding by way of demonstration, in this way the recipient helps the cardholder go on with her description. This is in line with Clark and Brennen (1991) principle of least joint effort: the notion that interlocutors are opportunistic and will select whichever method will take the least joint effort in order to communicate. This is in terms of time, resources and accuracy. Once the cardholder's difficulty is detected it is more time effective for the recipient to offer her interpretation for assessment than to let the cardholder exhaust further resources. Her assessment allows the cardholder an opportunity to confirm her interpretation as representative of the experience described or otherwise initiate her own repair (Clark and Brennen, 1991). In this case she confirms the validity of the assessment performed by the recipient both by nodding during the sigh and then verbally confirming in line 6: 'Yeah'.

The recipient's active involvement in describing the experience yields a mechanism that allows the cardholder to assess her recipient's interpretation, provide confirmation if she is correct, or if it were the case that her interpretation was inaccurate there is the opportunity to clarify and correct before she carries on. By performing her own interpretation of the action that results from that particular type of experience (muscle ache) performance of the embodied elaboration contributes as a function of intersubjectivity. As with the previous example, by giving an interpretation the recipient also avoids claiming direct epistemic access to the cardholder's recalled experience, respecting the cardholder's 'ownership' but still working towards empathetic communion (Heritage, 2011).

8.2 Non-congruent motor mimicry

As seen in Section 3.2.7.2 motor mimicry is defined as:

> An overt action by an observer that is appropriate to or mimetic of the situation of the other person rather than one's own (Bavelas et al., 1987, p. 317).

These displays are skilfully and precisely inserted into an interactive sequence to communicate an expression of how the recipient imagines or interprets the cardholder to be affected by the experience, displaying empathic affiliation with this response. As a cardholder's descriptions do

not always contain affect-laden expressions in some situations a recipient would produce a motor mimicry expression that is not a direct imitation of a cardholder's expression but an interpretation of how a recipient imagines a cardholder to respond to the experience. This is referred to here as non-congruent motor mimicry. The following examples show instances non-congruent motor mimicry expressions- these are not potential instances of automatic behavioural mimicry, as they are produced without an original expression to copy. The following discussion explores how recipients display empathetic understanding without a stimulus expression to copy.

8.2.1 Example 3: Surprise tokens in the absence of cardholder gesticulation

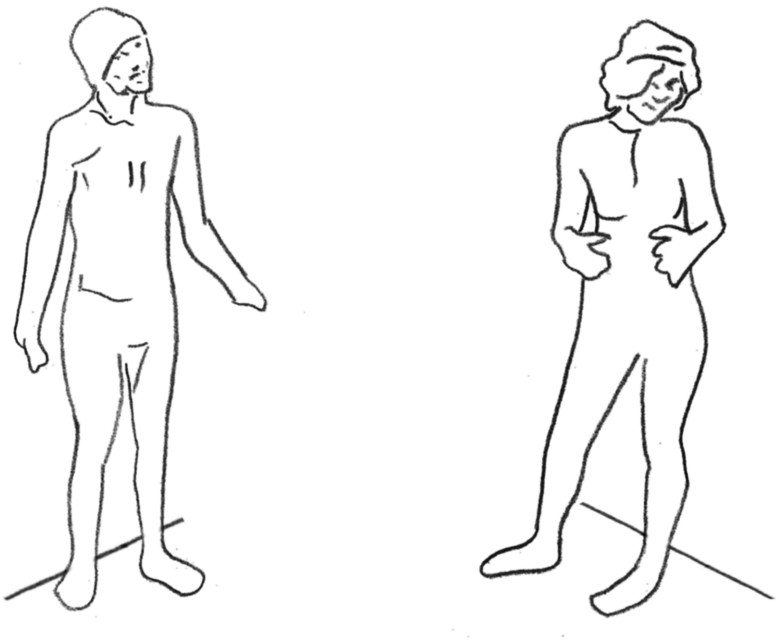
In this example the cardholder (In all figures in for example 1 the cardholder is positioned on the right of the sketch) is describing a method to pull out baby teeth as a child. His description provides contextual information; where he heard the method (at school), the apparatus used (a piece of string and a door knob) and acts out the action required (slamming a door) using iconic gestures. A descriptive gesture can be seen in Figure 8.10 where he performs the action of wrapping a piece of string around his tooth, gesturing towards the teeth deictically, another descriptive gesture can be seen in Figure 8.11 where he performs the action of slamming a door with his arm, making his hand into a fist as though he were gripping a doorknob. This action is accompanied by a verbal description of the action in lines 3 and 4. The cardholder doesn't produce any expressive reaction/pantomime of the pain it would cause, however his recipient does. Despite seemingly distracted in Figure 8.9 at line 1 where she is still considering previous talk (evident from her utterance in line 2), she attentively listens to the rest of the description from line 3. Immediately following the cardholder's performance of the door slamming action the recipient interjects claiming she knows the method from her own childhood. Figure 8.12 shows her quickly clasping her cheeks with her hands on the word 'kn::ow' at line 5 as if shocked from feeling the pain for herself. This isn't a simple case of motor mimicry in the traditional sense as the reaction could be construed as recall. Yet, her reaction is still appropriate to the function of demonstrating her understanding of the severity of the event the cardholder is describing. The recipient overtly renews her surprise at the method in playing the shocked recipient of the teeth pulling, an example similar to the surprise tokens as described in Wilkinson and Kitzinger (2014). The *surprise token* in this excerpt is an instance of what Wilkinson

and Kitzinger (2014) term 'ritualised disbelief' that does not signify 'genuine' surprise, but demonstrates an appropriate response- she is not reliving the surprise for herself but using its display as a communicative resource. Heath (2012) note that the interactional organisation of 'surprise tokens' in such contexts reflexively establishes an event or described event as legitimately surprisable (Heath, 2012). She performs surprise and a reaction befitting a sharp pain for the cardholder, embodying an emotive expression of the reaction expected in that particular situation. The expression forms an embodied 'news receipt' that functions as an 'interactional achievement' rather than a genuine expression of surprise and pain. The expression goes beyond solely providing a feedback response to affirm understanding but also provides evidence she is empathetic to the acutely negative nature of the cardholder's experience (Wilkinson and Kitzinger, 2014).

Goffman (1978) describes this surprise token as a 'centre of sympathetic attention', a type of *response cry* which is an exclamatory interjection defined as situational act that displays communicative alignment with on-going or told events such as expressions of revulsion, strain, pain, glee or in this case:- surprise. This type of display doesn't require acknowledgement and can be performed without interrupting the cardholder (Goffman, 1978). In this example however, the display accompanies the recipient's utterance 'I remember doing that with my parents' in Figure 8.12, and does interrupt the cardholder's description. She has already anticipated the closing action of the event so her talk overlaps with the cardholder's utterance 'and your tooth gets pulled out' at line 6. When she clasps her cheeks it is not a reproduction of his iconic gestures but a pantomime illustration of a response to those actions. This is not compatible with an automatic mimicry response as she does not mimic an expression performed by the cardholder but an appropriate response to the situation in the cardholder's perspective. She closes her response by saying 'it's horrible!' at line 7 and he responds by acknowledging her reaction- not offering sympathy back towards her expression- this would interrupt the flow of his description and as discussed in Goffman (1978) is not necessary. The example closes when he nods and laughs and then moves on the the next part of his story. The cardholder's laugh at the close of this excerpt is not necessarily because he finds her outburst funny but could be a punctuation effect acknowledging her response (Provine, 1996). The cardholder's positive acknowledgement of the recipient's contribution demonstrates to her that she has appropriately interpreted his description. The shared knowledge that they both understand the previous

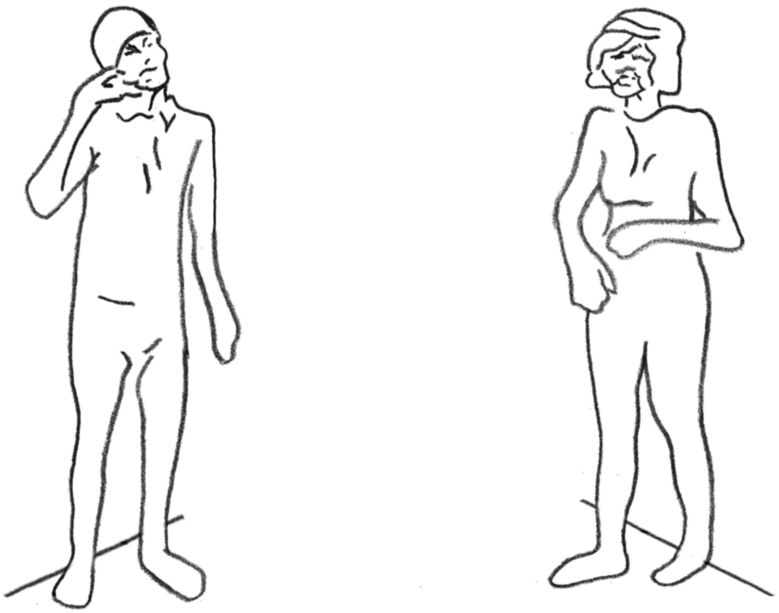
description allows further talk to be built upon and permits the conversation to progress.

Figure 8.9. Talk at school



1. CH: I- I'd heard [through school or something]
2. NCH: [.hhhh er you completely]

Figure 8.10. The tooth and the string



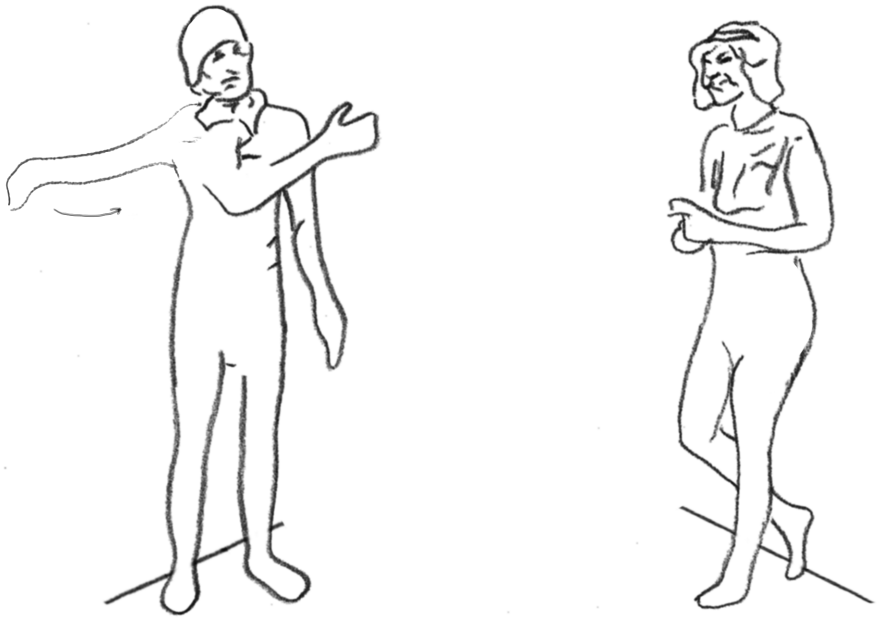
3. CH: you could tie a piece of string around it

(0.4)

4. CH: and then you attach the string to a door knob

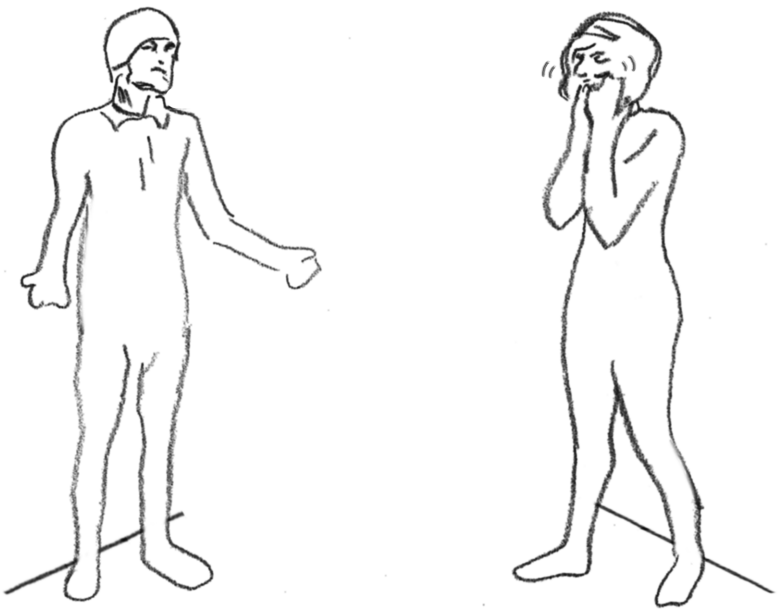
(0.4)

Figure 8.11. Slam the door



5. CH: and then you slam the door=

Figure 8.12. Pull out your tooth



5. NCH: =A:h yes I kno::w, [I remember doing that with my parents]

6. CH: [and your tooth gets pulled out]

(0.5)

7. NCH: it's horrible!

8.2.2 Example 4: Recipient performance demonstrating different aspects of experiences

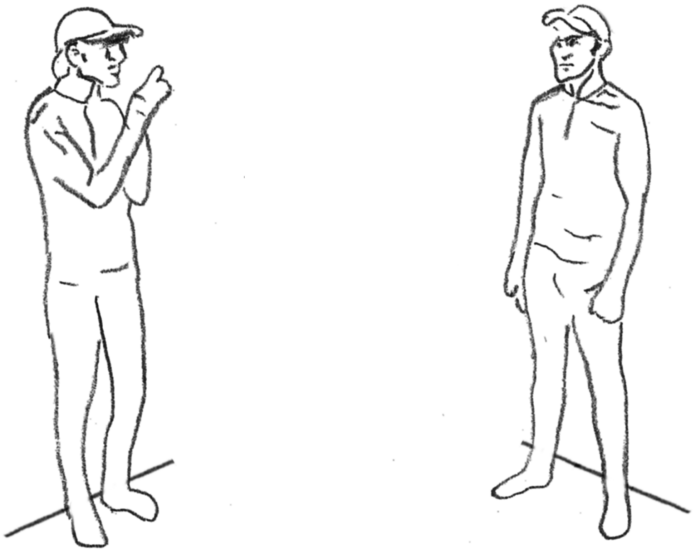
The following example shows another display of non-congruent motor mimicry. The cardholder (In all figures in for example 4 the cardholder is positioned on the left of the sketch) is describing a situation where one of his baby teeth is loose when eating. He begins in Figure 8.13 at line 1 'You know when your teeth were starting to fall out?' while he motions towards his mouth and producing a hand gesture that iconically represents the trajectory of a tooth falling out of his mouth into the gesture space in front of him at the same time as pointing at the left side of his mouth. The recipient responds by saying 'Yeah' at line 2. The cardholder immediately continues in line 3 'and one wouldn't qui::te' elongating the 'qui::te' while producing another iconic hand gesture that performs the action of pushing a tooth in the space in front of his mouth to demonstrate that the tooth was still attached but was possible to rock or wobble. He pauses for a moment and the recipient interjects in Figure 8.14 at line 5 with a 'Yeah', he has anticipated the rest of the cardholder's overlapping utterance 'come out' at line 4. The sketch shows the recipient pointing to his own mouth and grimacing while the cardholder carries on in line 4 'it was just that tiny bit'. The recipient offers an assessment at line 6 with 'it's horrific in retrospect' while continuing to grimace and motioning towards his own mouth with fingers gripped as though he is acting out attempting to tug out a tooth. The cardholder continues in Figure 8.15 at line 7 'you're sort eating or something', producing another iconic gesture of the act of using cutlery to scoop food into his mouth. Recipient continues to grimace while the cardholder says 'and you would bite down on it and it would just like' at line 8, on the word 'bite' the recipient produces an emotive expression of disgust with a gesture that could be interpreted as distancing himself, or creating a barrier between himself and the cardholder because he finds the scenarios repulsive and negative. His movements are exaggerated and dramatic, signalling that that his reaction is in response to distressing experience. The cardholder's speech pauses for a moment while he iconically gestures that his tooth is knocked out by pointing towards his mouth then slightly motioning his fingers forward out of his mouth towards the floor. He then says 'cr::k at line 9, to describe the sounds of the tooth tearing off the gums. To which the recipient responds with a grimace and says 'a::hhhh!' at line 10 and performs a biting gesture. The end of the recipient's 'a::hhhh!' at line 10 is slightly overlapped with line 11 in Figure 8.16 where the cardholder continues his description 'and it went cr::k and it was like a::h!'. The

cardholder continues to motion both hands towards the left side of his mouth slightly point outwards on the word 'cr::k', at this point he tilts his head in the direction he motioned his tooth to have fallen out before returning his hands back towards the left side of his mouth and straightening his head before he says 'and it was like a::h!' at line 12. At this point he switches from describing the sounds and motion of the tooth falling out and focuses on how it felt when he says 'a::h!' and exaggerates the tilting of his head to the left and holds the left side of his face in pain while grimacing. This forms a relatively subdued performance of his reaction to the incident at the time in comparison the recipient response. At line 13, the recipient joins in by saying 'a::hh!' overlapping the cardholder's 'a::h!' from line 12 where he grits his teeth in disgust then opens it again wide while pronouncing that 'a::hh!'. The cardholder smiles, relaxes his face and says '=and then it would fall out and there was blood everywhere' at line 14, while tilting he head to the left once more and motioning his hands from pointing towards the left side of his mouth to the left side of his waist, marking a fuller trajectory of the falling tooth. Returning to Figure 8.16, at this point the recipient bends over slightly to his right, he noticeably shivers, hunches his shoulders and shakes his hands in front of him. He then abruptly stops his grimaces and produces a short laugh at the point when the cardholder says 'and there was blood everywhere' at line 14 while smiling to mark the close of the description (Provine, 1996).

As can be observed in the sketches, the recipient's responses are more exaggerated than the cardholder's description and different in form. The cardholder mostly points to his mouth with deictic gestures and motions the trajectory of the falling tooth with iconic gestures that describe what caused the pain in detail, even reproducing the sound of the teeth snapping together and gums tearing when he say's 'cr::k'. The only point when he deviates from flatly describing the causes to enacting his reaction to the sensation when he says 'and it was like a::h!' by tilting his head towards the left, holding his face in pain and reproducing a groan 'a::h!'. The cardholder's expression is understated in comparison to his recipient, who writhes around in would-be pain throughout the sequence. Overall, the cardholder provides the narrative of the incident and the recipient performs a dramatic expression of the character of the experience.

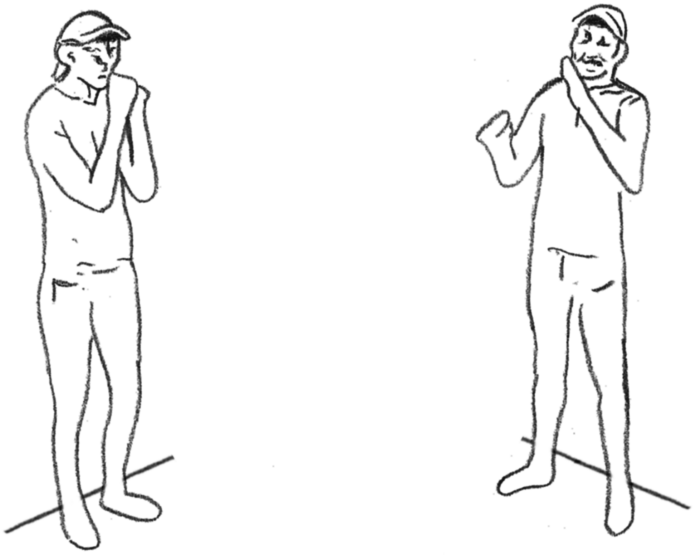
Like the previous example, a recipient demonstrating actions that express their interpretation of how the experience felt not only shows understanding but provides an opportunity for the cardholder to clarify when the recipient has seemed to misinterpret the description. When this

Figure 8.13. Baby teeth sketch



1. CH: You know when your your teeth were starting to fall out?
   (.)
2. NCH: Yeah
   (.)
3. CH: and one wouldn't qui::te
   (0.3)

Figure 8.14. Retrospectively horrific



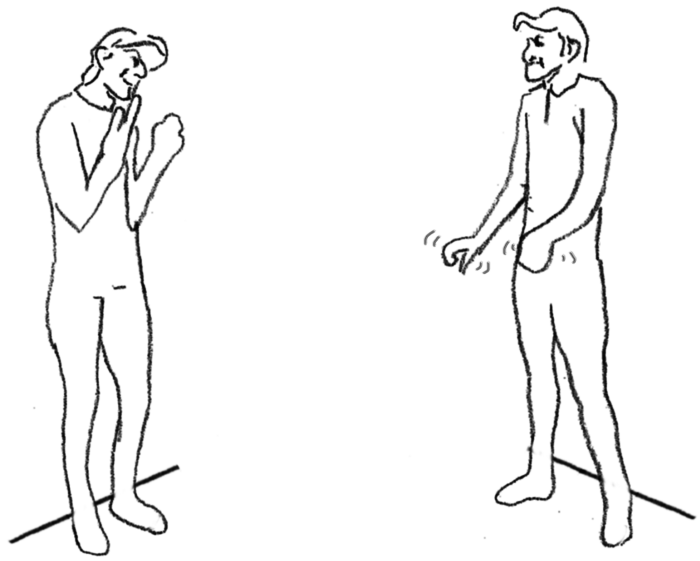
4. CH: [come out], it was just that tiny bit

5. NCH: [yeah]

(.)

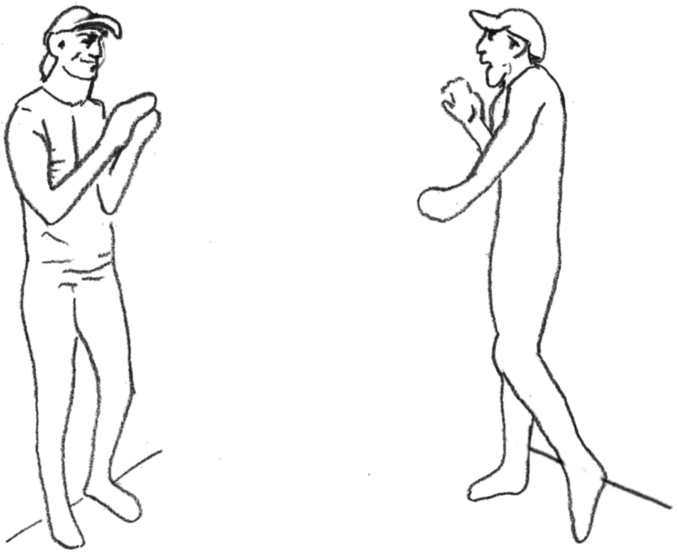
6. NCH: it's horrific in retrospect=

Figure 8.15. Bite down



7. CH: you're sort eating or something
   (0.3)
8. CH: and you would bite down on it and it would just like
   (0.3)
9. CH: cr::k
   (.)
10. NCH: a::h[hhh!]

Figure 8.16. Blood everywhere



12. CH: [and it went] cr:: and it was like a[::h]

13. NCH: [a::]h=

14. CH: =and then it would fall out and there was blood everywhere

is not the case all that's required of the cardholder would be to acknowledge his interpretation was accurate. Consider lines 9 through 13, the cardholder multi-modally describes the action of biting down on a tooth, the recipient grimaces along with the cardholder until he says 'cr::k' in line 9 when the recipient performs the action of biting himself. Sequentially the grimace then the performance of the bite occur at the appropriate time in the sequence, biting just after the 'cr::k' and then saying 'a::hhh!' as a response. At this point the cardholder could be said to providing an acknowledgement and acceptance of the recipient's performance in line 12 when he says 'and it went cr:: and it was like a::h' repeating the 'a::h' from the recipient's previous expression. The next utterance from the recipient is an additional 'a::hh!', to emphasise or comically reiterate that he understands the severity of the sensation.

Although they are both concurrently contributing to the description of the experience, there is no evidence of automatic behavioural mimicry (with the exception of the 'a::hh!' groan, however this is verbal) as each interlocutor is using different gestural strategies to demonstrate their understanding of the incident. As we can see, the description is collaboratively constructed and progresses sequentially, but most importantly intersubjectively.

These descriptions of experience are perhaps better understood as what Goodwin (1979) defines as joint interactive and concurrent constructions produced out of active collaboration.

Why is the recipient compelled to produce such an exaggerated and dramatic expression in response to the cardholder's comparatively subdued description of the experience? Wiggins (2013) argues that the enactment of disgust is an inherently social event. The recipient's lively enactment of disgust provides an appropriate response to the description of the toothache (Wiggins, 2013). The recipient's performance displays a reaction from the perspective of the cardholder's younger self as if it were his own (Bavelas et al., 1986, 1987, 1988). This both demonstrates his current perspective but also shows an intricate understanding of the experience being described, actively engaging in the in experience in a lively fashion.

It should also be noted that the cardholder specifically elicits a response from the recipient in line 1 'You know when your teeth were starting to fall out?'. A question not designed to request new information but to elicit feedback at the same time as interactively orientate the recipient to the event being described (Athanasiadou, 1991). This could be seen as what Heritage (2011) calls a 'pre-announcement' in a story-telling interaction, these are used by the story-teller to

project the type of action to be told and valence of the description. This gives cues as to the their stance on the story that facilitates an appropriate response from their recipient. This prefatory work solicits an empathetic response by inviting affirmation and affiliation with the tellers stance. The recipient's expressive response in this excerpt can be attributed in part to the cardholder eliciting an active engagement with his retelling.

8.3 Imitation for the purposes of clarification

The following two examples elaborate on the analysis on clarification questions and repair in the previous chapter. These previous analyses report a high rate of recipient movement when asking a clarification question and the cardholder also resorts to using their non-verbal resources more in these circumstances; usually a recipient is still when someone is speaking to them (Gullberg, 1998; Battersby, 2011). In line with this Beattie and Aboudan (1994); Duncan (1972) observe that gesturing places a claim over a speaking turn and acts as a interruption-suppression signal. By focussing on the sequence of gestures that the cardholder displays when clarifying ambiguities this analysis aims to investigate how embodied resources support intersubjectivity when there is a difficulty.

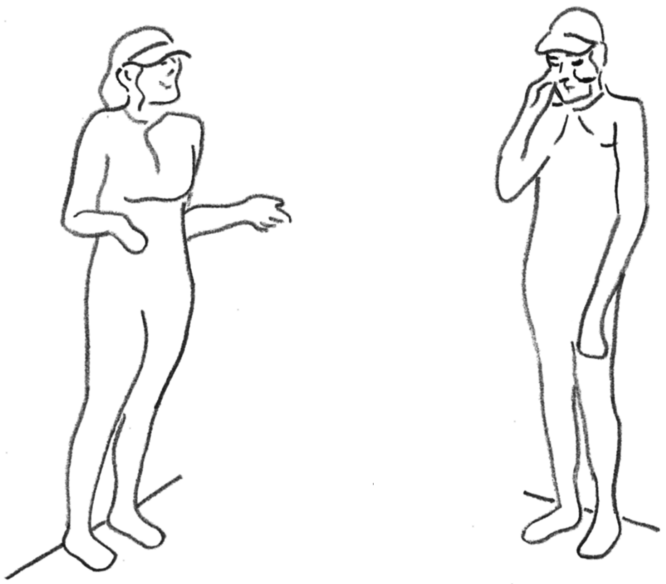
8.3.1 Example 5: Clarifying strange situations with movement

In the following example the cardholder (In all figures in for example 5 the cardholder is positioned on the left of the sketch) is describing an experience she had at a music festival in the UK. She explains that she enters a tent where there is an activity consisting of a laughing circle. In lines 1, 2 and 3 of Figure 8.17 she says 'and there was like this guy leading them', pause, 'in' another pause, 'a laughing circle'. This is said while she gestures with her arms out in front of her, palms facing upwards, in a typical shrug posture. Possibly pausing between utterances and shrugging could signify her own disbelief or amusement at the absurdity of the situation. The within-speech laughter over the utterance 'a laughing circle' serves as what Jefferson (1979) terms an 'invitation index' that a recipient can treat as providing a recognition point for recipient laughter. The recipient laughs dutifully at the end of line 3, producing feedback that acknowledges the comical aspect of the situation. The cardholder continues in Figure 8.18 to say in line 5 'and it was basically like everyone stood round, and he was like

doing like different laughs and stuff and it was like', pauses very momentarily to prepare for her pantomime of the belly laugh by raising her hands to her hips and pushing out her belly and verbalising the laugh by saying 'H:o! h:o! h:o! h:o!' in line 6, while bobbing her entire body up and down in time with each 'ho'. In Figure 8.19, at line 7 the recipient starts to ask 'Was it kind of like H:o?' over the second two of the cardholder's 'H:os' in line 6. He also raises his hands to his hips and pushes his belly out in a similar manner, but instead of bobbing up and down like the cardholder, the recipient leans back, sticking his belly out further and tilts his chin up in time with his 'h:o?', producing his own version of a belly laugh.

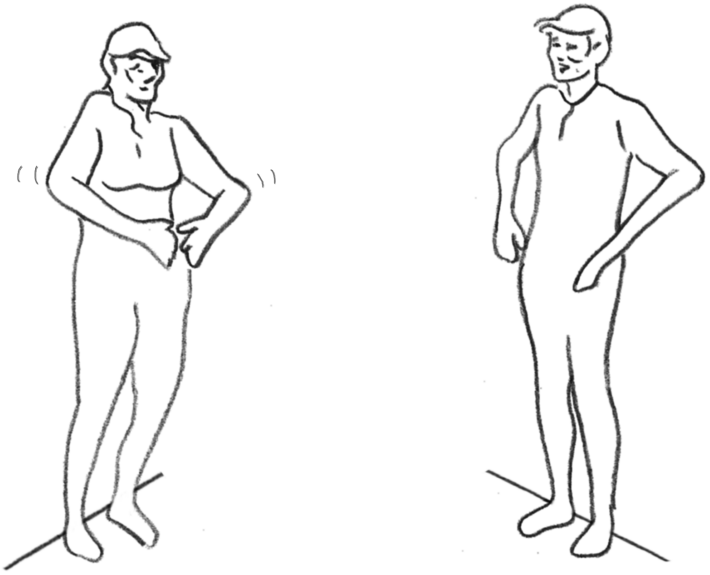
At this point the cardholder pauses her enactment in order to address the recipient's question; holding the gesture across his turn- an example of a post-stroke hold (Sikveland and Ogden, 2012). The cardholder also laughs in line 8 over the top of the recipient's version of the belly laugh pantomime, acknowledging that his enactment is also amusing to her. The clarification question asked by the recipient accompanied with a pantomime gesture is responded to in Figure 8.20 at line 9 'and it was, yeah and it was like, all like that', an affirmative acknowledgement that his enactment is accurate to the scenario she is describing. Line 9 sees the retracting her held gesture by gesturing interactively towards the recipient, signalling that the gesture he just performed is what she refers to when she says 'like that' at the end of her utterance, the retraction displaying resolution of the clarification (Sikveland and Ogden, 2012). These sort of expressions are what Bavelas et al. (1992, 1995) calls interactive gestures, this particular type functioning to cite other's contributions rather than as the cardholder producing a turn-selecting point signalling 'go ahead' (Battersby, 2011; Rieser, 2011). The recipient laughs at the point the cardholder motions towards him while referring to his expression, the interactive pointing gesture could also serve to elicit an affirmation the cardholder's story as an amusing anecdote, the recipient laugh showing accord with the cardholder (Provine, 1996). The cardholder carries on with her description after the end of the excerpt.

Figure 8.17. Laughing circle



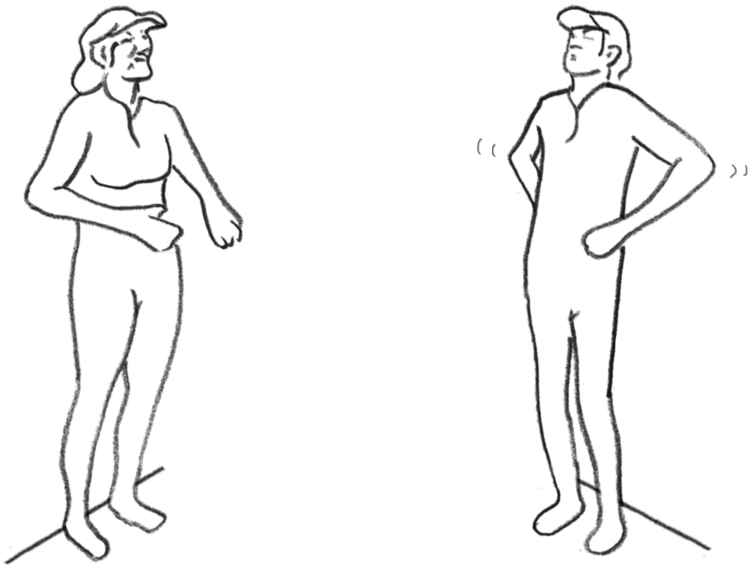
1. CH: and there was like this guy leading them
   (0.3)
2. CH: in
   (0.3)
3. CH: a laughing circle=
   (0.4)
4. NCH: =H:hh=

Figure 8.18. Ho ho ho



5. CH: =and it was basically like everyone stood round, and he was like
   doing like different laughs and stuff and it was like
   (.)
6. CH: H:o! h:o! [h:o! h:o!]

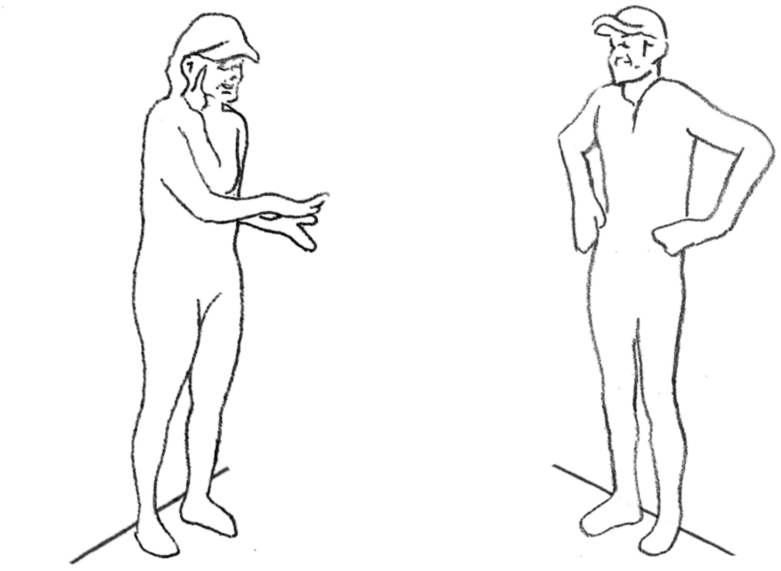
Figure 8.19. I go ‘ho ho ho’



7. NCH: Was it kind of like [H:o?]=

8. CH: [H:hhhh]

Figure 8.20. It was like that



9. CH: =and it was, yeah and it was like, all like that

On first glance, it could be said that the recipient was imitating the cardholder's enactment of a belly laugh when asking his clarification question. The situation the cardholder is describing is unusual and has already been marked with shrugs and pauses in the cardholder's opening turn (see Figure 8.17), it is hardly surprising that the recipient felt it useful to produce a demonstration as an opportunity for clarification. By suggesting that the situation was strange and out of the ordinary the cardholder could be said to invite the recipient to clarify as the situation might not be something he had experienced or heard of previously, so he would be introduced to the concept by her description alone. In the face of potential ambiguity he uses his embodied resources as an extra measure to defend intersubjectivity by clarifying actions in an unusual situations by repeating the gestures that are strange. The demonstration of his interpretation of the belly laugh gives the cardholder an additional opportunity to accept his interpretation or correct if necessary.

Similar to the expressive adaptions in the previously seen examples in Section 8.1 when producing stretching feedback along to the description of a yawn, the recipient uses variation and adaption in his clarification request to demonstrate his interpretation of the action. The cardholder can accept or correct him after one question, as opposed to him describing verbally a belly laugh. This would take longer and wouldn't add anything more than the action visually demonstrates. Although initially this sequence might be considered a direct copy, on closer inspection the two expressions are actually slightly different. As before, simply copying the action would not necessarily help the cardholder to determine whether the recipient understood the type of action she refers to, but would only show her he was watching her actions and holds the ability to replicate. Even in this case, the recipient would already possess an understanding of the experience to produce an adaption of it, which rules out the possibility that the recipient is imitating the cardholder in order to simulate the experience to aid understanding. Nor does it seem a product of an inner perception-behaviour link mechanism, as this would lead to automatic matching.

By using adaption and elaboration the recipient extends the action into that of a certain type, strategically allowing the cardholder both the opportunity to correct if he deviated outside of her description as she invites him to do so. Otherwise the clarification question serves to confirm the recipient understands and that the cardholder sees he understands. The cardholder affirmation in line 9 marks that his demonstration is correct and they are intersubjectively

aware that they are seemingly on the 'same page' and her description can move on with both of them engaged with the same experience.

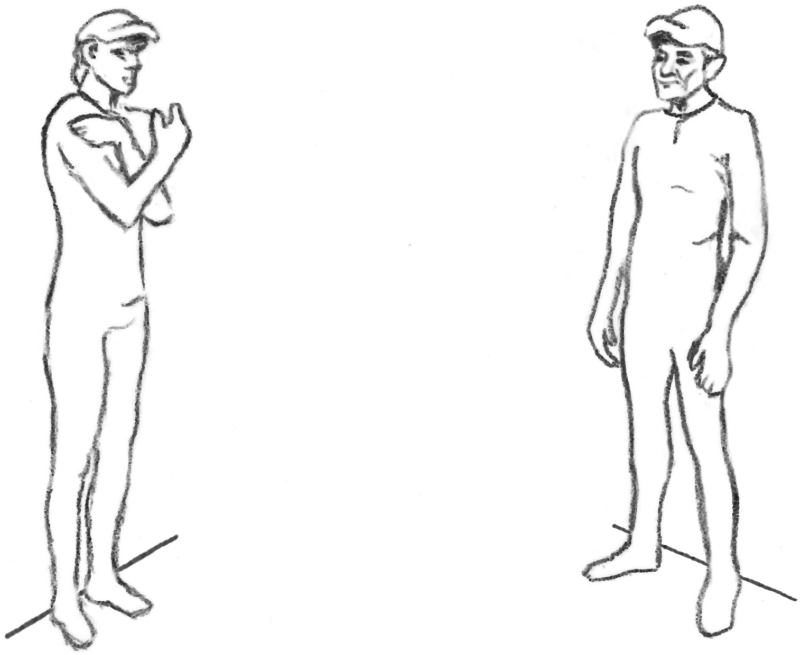
8.3.2 Example 6: Clarifying simple movements

In the following example the cardholder (In all figures in for example 6 the cardholder is positioned on the left of the sketch) is describing a technique to relieve the pain of a backache. Figure 8.21 shows the cardholder beginning his description in line 1 by asking the recipient 'have you had it where, if you've got a slight' pausing while he lifts his arms, starting to cross them in front of his chest with his hands moving to touch the opposite shoulder. He continues in line 2 by saying 'pain and people, were asked to do that?'. Quickly the recipient utters 'yeah' in line 3 in affirmation the the cardholder's question. As we have previously seen in example 4 (see Section 8.2.2) this participant uses the question form to confirm the initial context of the retelling is understood by the recipient (Athanasiadou, 1991).

The cardholder then continues by saying 'and people come up behind you and sort of lift- lift- lift you up' in line 4. During this utterance his hands start out on his opposite shoulders but when he says 'and people come up behind you and' he releases his hands forward slightly, then for the rest of the utterance he pushes his hands back onto opposite shoulder, seemingly going forward slightly to indicate force, then arches his back when he says the last 'lifts-', and tilts his head back to perform a backbend, then looks back at the recipient hand still on his shoulders completing his sentence 'lifts you up' with an upwards inflection on the word 'up' to indicate this utterance was still part of the original question. He pauses momentarily while looking at the recipient, waiting for feedback that he has understood. The recipient considers the movement but looks confused and says 'No' in line 5. The cardholder decides to change strategy and describe the outcome of the action he just performed by saying 'and that really cracked- that cracks your back' in line 6, self-repairing quickly to finish his utterance. At the same time as line 6, the cardholder releases his arms from his shoulders and points with his left hand to his lower back shown in Figure 8.22. The cardholder pauses for further feedback that the additional information might have clarified the action he intends to describe.

Still uncertain the recipient says 'what's that?'. Here there is a long pause of 0.7s, the recipient then adds 'you do that and someone?' performing the same enactment as the cardholder's

Figure 8.21. A slight pain



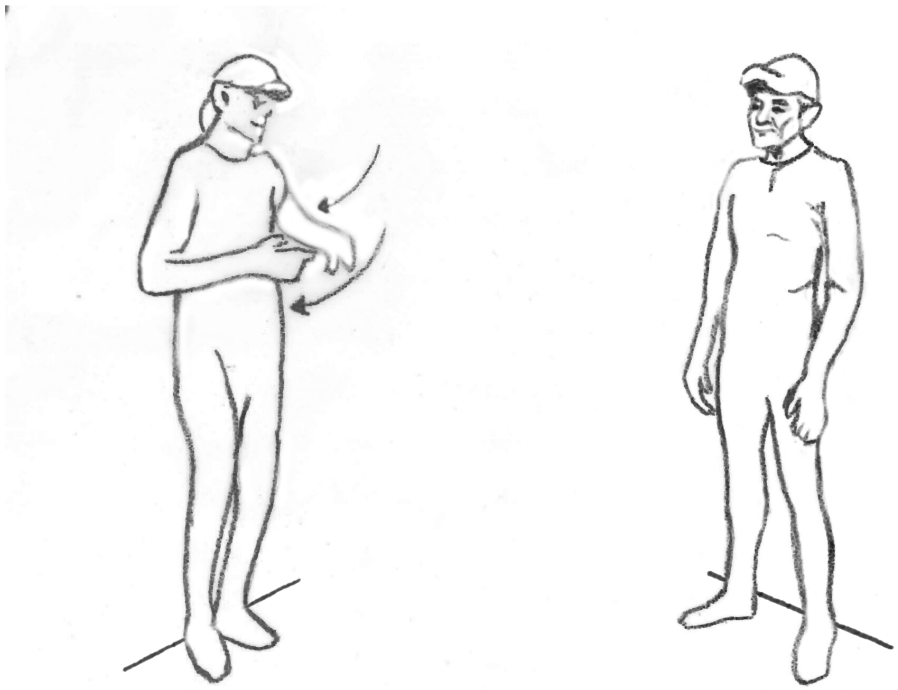
1. CH: Have you had it where if you got a slight
   (0.2)
2. CH: pain and people, were asked to do that=
3. NCH: =yeah

original (lifting his arms up and crossing them over his chest, each hand grabbing opposite shoulders). He continues to say 'and someone?' while leaning back into a back bend and tilting his chin up. The recipient's enactment and utterance in line 7 are performed in the same form, timing and intensity as the cardholder's original, displaying a straight imitation in Figure8.23. The recipient is cut short mid-question as his utterance overlaps with the cardholder's utterance in with line 8 that partly repeats his question with a clarification. Here the cardholder quickly interjects as the recipient has indicated what needed clarification, overlapping with the recipient's clarification question by repeating the utterance 'you do that and someone' placing his hands back on his shoulders. At this point the pair are exactly mirrored, both with their hands crossed on opposite shoulders. The recipient bends back first when he says 'and someone lifts you up' in line 7 which is immediately before the cardholder performs the same action while saying 'just like lifts you up from the shoulders there', patting his shoulders gently in line 8.

Clarification that the recipient had performed the correct action as well as the addition of 'from the shoulders there' leads the recipient to say 'oh no I've never heard of that' at line 9 in Figure 8.24, which serves as a confirmation that he understands the action despite not coming across it previously. Interestingly, the cardholder repeats the action for third time, but this time he adds a grimace, indicating the sort of reaction he would have if he experienced it. The repetition of the action is accompanied with the utterance 'and it clicks', then he moves his left arm to point at his lower left back again, but in contrast to line 5 when he says 'and that really cracked- that cracks your back' he also points over his right shoulder to his upper back while continuing to say 'like it clicks all the things down there and it's quite a nice sensation'. At the point of saying 'and it's quite a nice sensation', he moves he arms to his sides and wobbles them slightly as if to indicate the releasing sensation the action provides, that is the action of somebody pulling your shoulders back to crack your back. The recipient nods slightly at the end of the sequence.

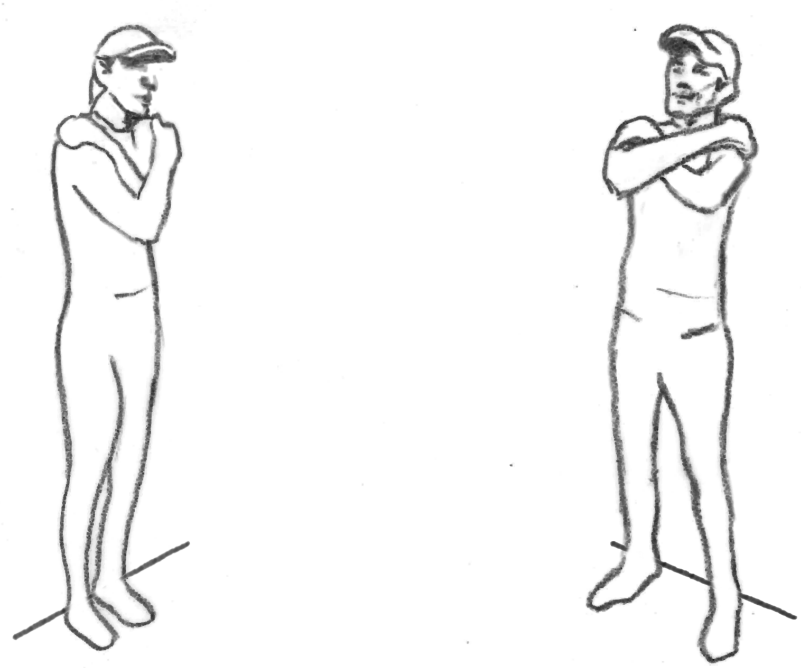
The action that is directly imitated, crossing the arms over the chest with hands on opposite shoulders and leaning back, is performed once by the cardholder, once by the recipient (the imitation), followed by two more repetitions by the cardholder. This example presents the clearest occurrence of straight imitation that occurs within 4 seconds of the original. So sequentially this could be an occurrence of automatic behavioural mimicry- facilitating affiliation and prosocial

Figure 8.22. Lifted up



4. CH: and people come up behind you and sort of lift- lift- lift you up?
   (0.3)
5. NCH: No
   (0.1)
6. CH: and that really cracked that cracks your back=

Figure 8.23. Somebody lifts you up?



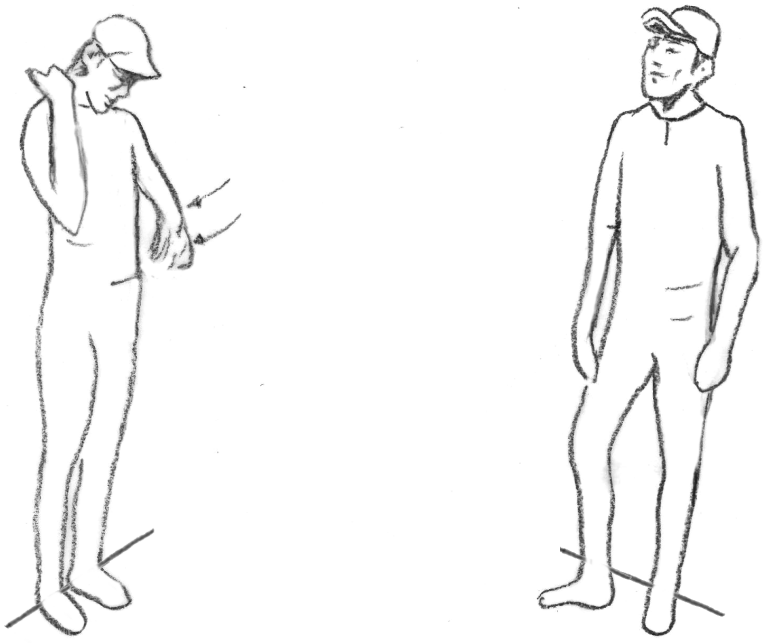
7. NCH: =what's that?

(0.7)

NCH: you do that [and someone?]

8. CH: [you do that] and someone just like lifts you up from the shoulders there

Figure 8.24. It clicks your spine



9. NCH: Oh no I've never heard of that

(0.1)

10. CH: and it clicks=

like it clicks all the things down there=

and it's quite a nice sensation

behaviour (Bailenson and Yee, 2005; van Baaren et al., 2009). However, the following analysis argues the mimicked actions are performed intentionally for the purposes of clarifying.

Line 7 presents the recipient's clarification question beginning 'what's that?'. After which there is a 0.7 second pause, as pauses between turns are typically 0.2 seconds (Zellner, 1994; Stivers et al., 2009) the recipient could judge that the cardholder is having difficulty disambiguating his talk (Healey et al., 2013) and upgrades his question with a stronger repair initiation that utilises additional nonverbal resources (the imitation) to help the cardholder locate the trouble-source action Schegloff (1997). The imitation is accompanied with a question 'you do that and someone?' in line 7, he times the imitation with 'you do that', as a component of a clarification question executed for the cardholder.

The cardholder confirms in line 7 'you do that and someone just like lifts you up from the shoulders there', here his reference 'that' refers to the imitated action. After the initial imitation the recipient performs, the cardholder performs it twice more. The cardholder's movements are precisely timed to coincide with 'you do that' so can be considered to be disambiguating his speech by performing. It appears that the re-enactment of the original is intended to re-confirm the action so the recipient knows he has performed the action accurately. The fourth and last performance of the action between them elaborates both with his speech in line 9 by saying while performing the action 'and it clicks' as well as the addition of a facial expression that describes additional information about how it feels. Although the action is the same, the additional details suggest that the last action is for the purposes of elaborating on the original action- this is in line with Holler and Wilkin (2011a) finding that recipient feedback that elicits detail or clarification causes cardholders to produce more communicative gestures in their storytelling with greater precision, size and visual prominence (Holler and Wilkin, 2011a). When the cardholder knows that the recipient has grasped that action he knows how to 'go on' (Mondada, 2011). Therefore when viewed in the structure of the interaction it appears that the imitation has precise and strategic interactional purposes.

8.4 Summary

All these examples are incidences of recipient nonverbal feedback expressions that are the closest probable instances of an automatic copying of body movement or instances of congruent

behaviour in the perspective of the cardholder.

The instances of what looks like mimicry in examples 1 and 2 contain expressions that adapt or elaborate on the original expressions, rather than straight repetition. These paraphrases are seen to be sequentially appropriate to demonstrate a richer understanding of the experience being described. Example 3 and 4 show instances of non-congruent motor mimicry, an expression that responds in a first person perspective to what is being described. Although these instances are not direct imitations of an interactional partner's expression, they embody the described experience. These displays respond to contextual details of the experience, such as pointing where the pain is. For instances such as this a recipient uses their embodied responses to demonstrate an understanding of the described experience by enacting appropriate behaviour in that perspective. Examples 5 and 6 are the closest to straight mimicry of these examples, a recipient performing a repetition of an expression performed by the cardholder. In each case the mimicked actions are performed intentionally and strategically for the purposes of clarifying observed movements rather than automatically copying perceived actions.

All these examples that look ostensively like examples of direct mimicry or repetition turn out to have richer communicative function when examined in the detailed interactional context. In particular, none of these cases involve automatic mimicry. Instead, some appear to suggest something analogous to verbal paraphrase serving to extend, clarify or elaborate on a prior turn. Taken together this suggests that people use embodied resources strategically within the context of particular interactional sequences where they provide useful additional resources for sustaining mutual understanding.

In conclusion, the examples analysed here exhibit evidence that nonverbal resources are intentionally and strategically placed as a function of higher level perspective taking cognitive processes parallels results from the previous three chapters. Chapter five finds the production of descriptive gesture are driven by conscious strategies interactionally involved in communicating a message or demonstrating understanding. The evidence found in chapter six finds that posture appears to be autonomous, or directed by constraints related to the act of communicating. Chapter seven finds that interlocutors are seen to adjust their nonverbal feedback strategies appropriately to the content, without directly imitating it.

Part V

Discussion

Chapter 9

Conclusion

9.1 Synopsis

This thesis assessed the role of embodiment for empathy when communicating about personal experiences. Given that inner experiences, such as bodily sensations and affective states, are not directly accessible to others; the question of how an experience is understood between people was explored. Building from the assumption that a common basic body plan provides a base level understanding of the perceptual world of Others, embodiment is regarded a key resource for empathy and intersubjectivity (Sinigaglia and Carrado, 2011). As the outward manifestations of these inner experiences are all that is perceptible to one another; this thesis considers how people demonstrate their understanding of another's experience through embodied interaction.

Accounts of intersubjectivity and empathy from cognitive neuroscience claim that by simulating the outward expressions of an inner experience, such as gestures, postures and facial expressions; embodiment is used as a vehicle to internally recreate an experience being described (Decety and Chaminade, 2003a; Singer and Lamm, 2009; Gallese, 2008; Chaminade et al., 2010). Research in social psychology and human interaction defines this process as an automatic *perception-behaviour link*, suggesting that it is this cognitive mechanism that facilitates automatic and nonconscious behavioural mimicry often observed in interaction (Chartrand and Bargh, 1999; Lakin et al., 2008; Chartrand and van Baaren, 2009; Chartrand and Lakin, 2013).

In contrast to prior research reporting an automatic and nonconscious propensity for people to mimic one another, the key finding of this thesis is that when talking about bodily experiences

the embodied responses used within these interactions are intentionally placed and formulated to facilitate the incremental process of conversation. This study has shown that people's body postures appear to become less similar when engaged in conversations about bodily experience. The types of gesture used in this context are dependent on the content or topic of the conversation and not reliant on the gestures of an interactional partner. Embodied feedback behaviours are constructed to demonstrate a specific understanding, or uncertainty in understanding, of the type of bodily experience being conveyed.

This thesis contributes to the above simulative theories of intersubjectivity by detailing how occurrences of behavioural mimicry, if any, are organised within a conversational structure, revealing an alternative explanation to the perception-behaviour link. Embodiment provides an important feedback resource for interaction not because it gives a level of similarity for understanding via simulation, but because it is integral to the collaborative, sequential and intersubjective progression of dialogue.

9.2 Summary of thesis

The hypothesis that connects the perception-behaviour link, the mirror neuron system and observations of behavioural mimicry in social behaviour, claims that cognitive simulation of another's actions and expressions are fundamental to intersubjectivity; automatic behavioural mimicry is a consequence of this process underlying this simulation. Numerous studies have found evidence of the ubiquity of behavioural mimicry in interaction, detailing marked prosocial effects such as empathy and rapport.

An issue with many of these previous studies that find empirical evidence of automatic behavioural mimicry within interaction rely on confederates, the problems of using confederates for experimental conditions have been highlighted in detail by Kuhlen and Brennan (2012). In these particular studies, confederates either mimic participants behaviour on purpose; or produce behaviours such as foot-tapping or face-touching that then appear in the participants behaviour (see Section 3.2.1). In either case the confederate produces behaviour that is not a spontaneous a component of the interaction, potentially effecting the natural course of interaction. The incidence of behavioural mimicry occurring between naive interlocutors engaged in natural, spontaneous conversation remains to be measured.

Moreover, these studies analyse mimicry of non-content specific mannerisms such as foot-shaking and/or face-touching that are not performed as a means to communicate the content of talk. For this reason, the inner simulation of these mimicked behaviours would not lead to an insight to what is being spoken about, so cannot contribute to intersubjectivity or empathy in this context. An embodied inner simulation of the expressions of another is claimed to contribute to an attribution of mental states to one another. Mimicking foot-shaking and / or face-touching does not allow the opportunity for this to occur. Additionally, the interactional situation (i.e. describing photographs) is not conducive to provoking empathy between people. Nor do these interactional situations in the previous studies assess the impact of valence on behavioural mimicry, for example describing positive or negative experiences.

Existing studies that find posture congruence between interacting parties positively correlate the imitation of postural configurations with rapport and empathy. However, the limitation for many of these studies is experimental factors that provide a lossy and approximated account of postural activity. For example, using broad classes of predefined body configurations, sampling at large time intervals and / or focussing on a small range of body parts. A granular, full-body and non-binary measure of postural similarity over a continuous time-series would reveal a more conclusive account of posture congruence of people engaged in empathetic interactions.

The global nature of these studies- i.e. measuring the frequency of certain behaviours over the entire course of the interaction does not provide any insight into the organisation of mimicry within a conversational sequence. Several studies that observe gestural mimicry within interaction (see Section 3.2.7) suggest that the recycling of gestures are strategically utilised to facilitate forming a mutually shared referent expression, representational gestures being repeated to claim, establish and display understanding (Kimbara, 2008; Mol et al., 2012; Parrill and Kimbara, 2006; Furuyama, 2000). This is in line with the general hypothesis that representational gestures are performed intentionally for the benefit of an interactional partner (Alibali, 2001). Again, these studies focus on interactional situations outside of empathetic exchanges, whether mimicked gestures that describe personal experiences are used in a similar way, considering that the experience they describe is private and inaccessible between people, is yet to be investigated. Does a closer look at instances of behavioural mimicry in an empathetic interactional setting reveal an alternative to the perception-behaviour link as its basis?

Models of recipient feedback claim that these signals are crucial for intersubjectivity as they

provide an immediate and incremental indication of understanding or uncertainty in the course of an interaction (Heylen et al., 2011; Ward and Tsukahara, 2000; Allwood et al., 2007; Allwood, 2003; Cerrato, 2002). The question of how behavioural mimicry integrates into accounts describing how people demonstrate their understanding (or problems in understanding) through feedback behaviours remains to be investigated. Specifically addressing how embodied resources are used to give empathetic feedback as deployed as a sequence of turns (as observed in (Heritage, 2011; Kupetz, 2014)).

9.3 Research questions

In order to contribute research filling the gaps identified above, three broad research questions were posed:

1. When engaging with one another's experiences, do interlocutors exhibit automatic behavioural mimicry of one another's content-specific gestures, postures and expressions in spontaneous dyadic interactions?

2. Are conversational resources regarded as important for intersubjectivity, such as feedback expressions and repair, used strategically when working towards empathetic understanding?

3. How do instances of mimicked gestures, postures and expressions fit into the sequential organisation of feedback expressions and repair in this context?

Using a combination of quantitative and qualitative methodology to investigate the contribution of nonverbal resources to empathy and intersubjectivity, chapters five, six, seven and eight analyse a corpus of dyadic interactions featuring recalled embodied experiences, such as painful or pleasant bodily experiences. This was compared to a separate corpus of dyadic interactions featuring informal conversations about current affairs to investigate how empathetic content influences interlocutors use of embodied resources, the analysis tested whether interlocutors converge in gesture use and posture or not, and what embodied feedback mechanisms are used to demonstrate empathetic understanding.

9.4 Chapter summaries

The first research question is addressed in chapters five and six. These chapters measure the incidence of gestural mimicry and postural congruence within interactions between two participants with a goal to share and understand one another's descriptions of recalled bodily experiences, a situation that is expected to promote empathetic interactions experience.

Chapter five measures the patterns of gesture type and duration as observed from annotations from audio-visual analyses of both corpora. The findings show that participants in cardholder and recipient roles use different patterns of gesture type and duration. Although the overall pattern of use for iconic, deictic, pantomime and metaphoric gesture types is not reliably different between roles, the use of abstract descriptive gesture differs significantly. Recipients appear to focus more on the contextual information than the description of sensation of the experience itself, in line with Heritage (2011) observation that the epistemic rights to an experience are asymmetrical between the person describing the experience and the audience of the description. Again consistent with observations from Heritage (2011), recipients also appear to harbour a stronger moral obligation to engage with descriptions of negative experiences, by producing longer gesture durations overall. Patterns of gesture type and duration are also determined by the content of the interaction as it differs between corpora with different topics, indicating that different gestures are useful for communicating about different types of conversations. This is important as it questions the validity of the claim that expressions are automatically copied as predicted in Section 3.2.1, rather gestural patterns appear to be tailored to the content of the talk, the audience of the gesture, and the role the producer of the gesture has within the interactional context.

Chapter six analysis uses a more computational approach to detect similarity between interlocutor's expressions by comparing their morphological form via motion capture analyses. The body movement data collected allows a quantitative test to determine the degree of similarity between participant's body postures found within these interactions. The findings tend to reflect those found in chapter five; variation in posture differs for descriptions of different experience types, for example cardholders vary their posture more for toothaches and yawns than descriptions of other experiences, seemingly because for these topics interlocutors are more likely to raise their hands to their mouth; to point at a location of a toothache or stretch out

to describe a yawn. But this is not influenced by the interaction. Also, recipients differ from cardholders in that they tend to hold the same posture for extended periods. Overall, the average posture similarity between interlocutors is not reliably different from expected chance levels. There is no evidence of automatic mimicry sustained over the course of an interaction in the posture similarity data.

The second research question is addressed in chapter seven investigating the ways in which interacting parties acknowledge and engage empathetically with one another's descriptions of bodily experience via nonverbal feedback responses. Following Heritage (2011) and (Kupetz, 2014), recipients and cardholders are predicted to organise their feedback responses appropriately to descriptions of personal experiences, interlocutors will respond and engage with these accounts in a strategically unique manner than to other types of dialogues. This analysis measured organisation of feedback signals coded from audio-visual annotations from both corpora. The results show that recipients produce the majority of feedback signals and the function of these signals is appropriately empathetic to a cardholder's description. Reciprocally cardholders produce feedback appropriate to their role. Feedback strategies found in interactions based around accounts of bodily experience differed significantly from interactions based around informal conversation about current affairs in the control dyads. Recipients adjust their feedback strategy and use their embodiment in ways appropriate to the material, without replicating the expressions of the cardholder. The hypothesis recipients would respond more strongly to accounts of negative experiences due to an expected tendency towards altruism and sympathy (Preston and de Waal, 2002) as interlocutors did not produce more nonverbal affective feedback responses; there is no influence of valence on the pattern of feedback for either cardholder or recipient. The pattern of empathetic responses appear to follow the same interactional strategy whether the topic is positive or negative.

The analysis in chapter seven also examined the embodied resources used during the process of repair, specifically where recipients flag a point of understanding that needs to be addressed before the conversation can proceed by directly requesting clarification. Nonverbal communication is considered useful at these points; if an interactional resource is not suited or insufficient interlocutors are expected to turn to alternative or multi-modal resources to compensate. This analysis measures the production of hand gestures during repair sequences using body movement data and coded gestures from audio-visual annotations. The findings indicate that there is

a higher level of collaboration at these points where both parties contribute more nonverbally to work towards resolving the problem. Outside of clarification sequences, non-speaking addressees do not move their hands as quickly and tend not to produce as many descriptive gesture as speaking participants. They do provide concurrent feedback to signal continued attention and understanding. However, during clarification sequences, the participant currently speaking and the non-speaking addressee tend to merge their nonverbal behaviour. Non-speaking addressees producing more than twice as many descriptive gesture in overlap with a speaker's turn at these points than otherwise. This underlines the collaborative nature of conversation, the strategic importance of nonverbal resources for sustaining mutual understanding and the critical role of feedback and repair in working towards intersubjectivity.

The third research question is addressed in chapter eight. The excerpts analysed in this chapter are selected as the closest examples of automatic mimicry present in the corpus. From analysing how the behaviour observed as mimicry in these examples are organised around the structure of interaction, the behaviour can be seen to serve strategic communicative function for the purpose of collaboratively working toward intersubjectivity. The instances of what looks like mimicry in examples 1 and 2 contain expressions that adapt or elaborate on the original expressions, rather than straight repetition. These paraphrases are seen to be sequentially appropriate to demonstrate a richer understanding of the experience being described. Example 3 and 4 show instances of non-congruent motor mimicry, an expression that responds in a first person perspective to what is being described. Although these instances are not direct imitations of an interactional partner's expression, they embody the described experience. These displays respond to contextual details of the experience, such as pointing where the pain is. For instances such as this a recipient uses their embodied responses to demonstrate an understanding of the described experience by enacting appropriate behaviour in that perspective. Examples 5 and 6 are the closest to straight mimicry of these examples, a recipient performing a repetition of an expression performed by the cardholder. In each case the mimicked actions are performed intentionally and strategically for the purposes of clarifying observed movements rather than automatically copying perceived actions.

9.5 Contributions to knowledge

Taken together, these findings offer some important contributions to the study of embodiment as a resource for intersubjectivity and empathy:

1. It has been demonstrated, using a combination of quantitative methodology, that interlocutors do not automatically mimic one another's descriptive gesture or overall postures in interactions about bodily experience:

   (a) A full bodied and highly sampled analysis of posture congruence via body movement data shows posture similarity between interactional pairs is not above chance levels, posture is not automatically matched to an interactional partner's.

   (b) In addition, it is shown that patterns of content-specific gesture production differ between interactional partners.

2. Alternative explanations accounting for observations of mimicry have been proposed. Using qualitative methodology examining a selection of examples of what appear to be mimicry in a corpus of dyadic interactions based on descriptions of bodily experience, various communicative functions behind these instances are presented. These demonstrate that interlocutors use nonverbal feedback signals strategically to supply evidence of empathetic understanding toward descriptions of bodily experience.

3. Hand-coded annotations of instances of feedback have been aligned with accounts of empathetic response patterns as presented by Heritage (2011) and Kupetz (2014). Providing quantitative evidence to support their observations of when people work towards empathetic communion.

4. Further evidence that interlocutors use their embodied resources according to the type of content being discussed is shown. People use characteristic movements to describe particular bodily experiences, as well as using differing patterns of gesture type use for interactions in different interactional contexts. Differing interactional contexts have also been shown to influence the type of nonverbal feedback responses produced.

5. Evidence has been provided that suggests that a narrative structure and the roles determined therein has an influence on how interacting parties coordinate their nonverbal

behaviours. This is in terms of posture variability, gesture production and feedback signals.

6. It has been demonstrated that interlocutors use a different organisation of embodied interactional resources during clarification sequences than otherwise, such as a different pattern of gesture production and feedback signals across turns.

These contributions are important because they allow us to account for the sequential organisation of embodied resources by which interlocutors can work towards intersubjectivity. By integrating analysis on how different modalities feature when engaging empathetically with one another, this research gives a wider view of the structure of communicative mechanisms underlying the ongoing process of interaction, contributing to building embodied conversation agents and human-computer interfaces. The implication of this work on research on automatic behavioural mimicry provides an insight into a body of work that describes the role embodied interactional devices play in affiliative and prosocial behaviour, advancing research in the fields of human interaction and social psychology. The work also sheds light on the behavioural counterpart to the proposed cognitive architecture detailed to simulate one anothers experience by cognitive neuroscientists. By conceptualising empathy within practises of interaction this research adds empirical evidence to the wider phenomenological question of how we come to understand the experience of the Other.

9.6 Implications of contributions on previous research in the field

9.6.1 Evidence of behavioural mimicry

A wealth of research has referred to evidence of the ubiquity of behavioural mimicry between people during interaction, often highlighting prosocial consequences of this tendency on empathy and rapport. Section 3.2.1 reviews literature that claims behavioural mimicry is regarded an automatic response to perceiving the actions of others, caused by a perception-behaviour link (Chartrand and Lakin, 2013). Research on the mirror neuron system describes neuro-physical evidence behind a perception-behaviour link (Iacoboni et al., 1999; Decety and Chaminade, 2003b; Decety and Sommerville, 2003; Wang and Hamilton, 2012). These accounts claim that an

inward simulation the manifest expressions perceived lays the groundwork to an intersubjective understanding; these mechanisms are what leads to an automatic propensity for behavioural mimicry (Chartrand and Bargh, 1999; Yabar et al., 2006; Stel and Vonk, 2010; Lakin and Chartrand, 2003; Lakin et al., 2008).

This work measures the incidence of behavioural mimicry in interactional context that promotes empathetic engagement between interlocutors. As these interactions feature discussions of bodily experiences that, based on a shared experience of being embodied, can be empathetically understood; the link from imitation to simulation is proposed to contribute to the experience of empathy between people. As these findings did not observe any mimicry in posture or descriptive gesture, this work does not demonstrate evidence supporting this model or the perception-behaviour link.

The majority of these previous studies that make the above connection between mimicry behaviour and simulation rely on instances of mimicry where the behaviour measured (foot-shaking / face-touching) does not directly contribute to the communication of a message (although perhaps an indication of the stance or attitude of the interlocutor). The present study is not directly compatible to these results as it observes descriptive gestures that do contribute to a communicating a message. However, it does not verify this hypothesis behind this connection as it cannot supply evidence that extend these simulative theories into behaviour that conveys semantic content. Intuitively, behaviour that conveys the content of talk can be directly simulated in this manner.

The data provided here that shows the levels of posture congruence are not above chance levels is in contradiction with studies that report behavioural mimicry is ubiquitous in interaction. Although the context (interactions with confederates vs spontaneous dyadic interactions) and method (annotated behaviours from video vs body movement data) that the data collected here reports on differs from the above studies, the presence of a general prevalence of behavioural mimicry would be expected to be appear in a systematic approach to measuring body movement similarity. At the least, mimicry is not as automatic as claimed above in all interactional contexts.

Previous studies that directly measure posture congruence in a range of interactional contexts claim a high level of posture congruence between people, but specifically state that the higher

the level of congruence induces a higher level of rapport (Scheflen, 1964, 1972; LaFrance, 1979; Trout and Rosenfeld, 1980; Maurer and Tindall, 1983; Feese et al., 2011; Tia et al., 2011; Hagad et al., 2011). The present study does not measure rapport but does demonstrate that the level of posture congruence between people is, on the whole, very high simply by virtue of the physical constraints relating to a common body plan, not necessarily due to underlying copying of postures. In addition, the present study uses a finer and more granular measurement than those used in these previous studies that claim a link between posture congruence and rapport, or that posture congruence is an index of shared attention and comprehension.

9.6.2 Mimicry as communicative

A closer moment-by-moment examination of a small sample of excerpts from the bodily experience corpus were thought the closest probable instances of automatic mimicry or mimicry behaviour, look directly for a communicative function within the sequential organisation of the unfolding interaction. The qualitative report that comes from this analysis highlights some interactional strategies behind the mimicry behaviour reflected on. These strategies align with those attributed to reports of motor mimicry (see Section 3.2.7.2) and gestural mimicry (see Section 3.2.7.1). Demonstrating further corroborative accounts of the way embodied resources are exploited when interlocutors work towards intersubjectivity.

The first and second example presented in chapter seven describe sequences that contain expressions that adapt or paraphrase an original descriptive expression to demonstrate understanding of the type of experience being described. These instances extend research on gestural mimicry that reports the rephrasing and recycling gestures is used to show understanding and agreement in a learning environment by highlighting similar occurrences in a different interactional context (Tabensky, 2001; Furuyama, 2000). A number of studies have delved further into gestural mimicry of this function, these studies contain observations of interlocutors adapting one another's gestures to converge with previous gestural contributions, suggesting that these mimicry behaviours form mutually shared referent expressions or conceptual pacts (Mol and Krahmer, 2009; Mol et al., 2012; Kimbara, 2006; Holler and Wilkin, 2011b). This work adds an alternative interactional scenario that this strategy is used in, showing that when converging gestures people not only establish known expressions that refer to something concrete that is mutually understood, but also exhibit an understanding of one another's experience through

adapting or personalising the outward expressive performance of it.

The third and fourth example describe sequences that contain non-congruent motor mimicry after Bavelas et al.'s 1986; 1987; 1988 definition. The examples extend this research by providing further evidence that motor mimicry expressions have a communicative function. Expanding the definition to include use in circumstances where the expression is used to show empathetic understanding toward descriptions of bodily experience.

9.6.3 Empathetic feedback and repair

Numerous studies have presented models of feedback functionality, highlighting the general view that feedback enhances the fluency and effectiveness of interaction (Bavelas and Chovil, 2000; Bavelas et al., 2006; Allwood et al., 1993; Cerrato, 2004, 2007). Models of feedback have offered explanations to the function of specific feedback types typically found in interaction, such as backchannels and affect bursts. Using a coding scheme derived by this body of research a pattern of feedback functions was determined to compare the corpora studied in this thesis. This places a generic model of feedback into a wider interactional organisation to contribute to defining the interactional characteristics of empathetic dialogue, adding further verification of the observations of empathetic response strategies made by Heritage (2011); Kupetz (2014).

The data presented on clarification sequences in this thesis parallels with Holler and Wilkin (2011a) work on gesture during clarification requests or confirmatory feedback. Their study shows a similar influence on gesture production over the different stages of a clarification sequence. Together, these studies align to demonstrate that embodied resources are an important modality in working towards intersubjectivity at points of juncture such as repair.

9.7 Critical assessment of the thesis

There are various strengths and limitations found within this study that must be taken into account when considering the contributions presented here. These are detailed below.

Working with motion capture data

The body movement data collected in this work used the motion capture system described in

Section 4.4 had the advantage of providing precise three dimensional measurements of each participant's movements accurate to the millimetre, and fine-grained enough to provide data points at up to 120 frames per second. This allowed for a systematic and quantitative analysis that is not vulnerable to disadvantages of slower labour-intensive video-based techniques, such as the frame-by-frame hand coding conducted by Condon and Ogston (1966) which is susceptible to human error as well as annotator inconsistencies (Condon, W. S. & Ogston, 1966). In comparison, motion capture data provides a less time consuming but more precise data set for analysis. The advantage of using motion-capture over cheaper computer-vision human movement tracking techniques such as surveyed in Gavrila (1999) and Poppe (2010), is that motion capture retains an accurate representation of the three dimensional element of movement; which current computer-vision techniques only approximate.

The disadvantage of using motion capture to record body movements is that participants are required to wear lycra motion capture suits in a lab setting. Although the suit does not hinder a participants movement, it is uncertain whether this might affect the interaction. Another issue is marker drop-outs (see Section 4.4.0.4 for a more detailed description), which can result in missing data points.

Movement data in itself it does not describe anything meaningful. The challenge is to link characteristics in the numerical values with meaningful interactional events. For example, a known gesture such a thumbs up sign is readily recognised by a human observer but to be detectable from motion capture values an algorithmic process describing it's particular physical properties would need to be determined. The advantage of this type of analysis is that the parameters that define these events will always remain consistent. However, the problem with explicitly defining these properties as *events* within an interaction is that the process will also detect movements that fall within the predefined computational parameters but are not actual instances of the interactional event. As we have seen the drawback of this type of analysis does not take account of the context of the talk under analysis, so the results only provide an index of the interactional event they define. An index can still be a useful tool if this is taken into account when drawing conclusions from the results.

The posture congruence analysis presented in chapter six presented a measure of self similarity that compared pairs of joints between two interlocutors. This meant that the right elbow of one participant is compared to the right elbow of another participant. This confiuration

assumes that mimicry is defined under a carbon copy distinction as opposed to a mirror image distinction (Bavelas et al., 2002). There is no concensus as to which distinction defines posture congruence in the existing literature. This difference could pose a difference in similarity results when participants posture has large variations between the left and right sides of the body. In addition, the posture congruence analysis here presented an overall measure, including a full body analysis. This does not consider that posture congruence could be specific to particular body parts. For example, similarity could only appear for the upper body, or the trunk of the body without the arms. Only considering an overall full body measure could mask any significant posture similarity above chance that might occur in distinct configurations of body parts.

Issues with hand-coded annotation and qualitative analysis from video

Whilst human analysis of the video is contextually richer than automated processes, using human coders to interpret data accentuates the problem of *experimenter bias*. There are several problems analysing video data. Although in general participants habituate to the situation of the lab and absorb themselves in the task of interacting, there is a possibility that participants behaviour is altered as they are aware the camera is watching them. To mitigate this the first two items were disregarded (see Section 4.2.4) an initial settling in period, but it is uncertain as to the prevailing effects of this issue.

Camera angle could also influence the interpretation of interactional events, as it does not capture the same perspective as the participants. This is generally more of a problem in interactional settings that require approximating spatial factors in their analysis, but has less impact on the interpretation of gesture taxonomy, feedback function, identifying clarification sequences or overall structure of nonverbal behaviour that this thesis addresses.

These issues have a broader influence on qualitative analysis in chapter eight which can also be subjected to experimenter bias. This problems occurs when an experimenter overlooks or focusses on features that support a prior hypothesis or intuition that is not based on objective observations in the same way as statistical analysis. Qualitative analysis also bears the limitation that there is no guarantee that the examples selected are demonstrative of normative behaviour.

Statistical results based on annotations of gesture type and feedback function have the limitation

that they are temporally global, i.e. the annotions occur at any point during the interaction. This means the analysis of the feedback functionality lacks a sequential element and cannot comment on the sorts of sequential organisation that Kupetz (2014) reports. This limitation is a particular issue for the analysis of gesture type production when investigating for instances of gestural mimicry. A closer inspection of gestural overlap, or an inspection of which gesture mimics which would shed a broader view on which participant (cardholder or recipient) typically tends to mimic, and if this occurs in a common set contexts. As previously noted, quantitative analysis also presents a global insight in terms of gesture type and it is unknown whether individual gestures are copies just by virtue of being of the same gesture type.

Limitations of sample base

As the contributions that this work produces are based upon the corpora analysed, claims beyond this cannot be reliable stated. Therefore the results presented here are only demonstrative of the sample base analysed and does not account for all conversations. For this reason, limitations in the scope of the sample base must be considered as they limit the scope of the contributions posed.

The bodily experience corpus analysed ten interactional pairs and the control dyads also analysed an additional ten interactional pairs. This sample base is minimal, although enough to form significant statistical observations, any atypical behaviour occurring in the corpora has the potential to skew the results further than a larger sample base as there is less data to smooth out statistical anomalies.

All participants were fluent English speakers, however of a small number (not accounting for more than a quarter) English was not their native language. As mentioned previously, gesture production increases with the sense there are problems with understanding or when communication is difficult (Bavelas et al., 2002; Holler and Beattie, 2003). This means there is the possibility that gesture production in interactional pairs where one or both participants were non-native English speakers increases due to this factor.

In the bodily experience corpus, half the pairs were strangers and the other half previously knew each other as friends (the control dyads only contained participants who did not previously know one another). This presents a factor of familiarity to the interactions between participants who previously who one another, prior knowledge of the content of a description could cause

recipients to behave differently than if the description was fresh. This could especially effect empathetic responses as a recipient may have already demonstrated their understanding to parts or all aspects of the description of the experience.

9.8 Future research

9.8.1 Outside the dyad

The findings shown in this thesis highlight the asymmetrical pattern in the use of embodied resources between interactional roles in a narrative context. Recipients are seen to align these resources to demonstrate their understanding strategically where they have been observed to use elaborations or adaptions of a cardholder's expression. The corpus analysed here considers only dyadic interactions, this raises the question of how this mechanism translates to a multiparty context? How is intersubjectivity collaboratively attained past a dyad?

- Is intersubjectivity regarded on an individual basis, where each recipient produces their own expressions to demonstrate their own understanding that is unrelated to another recipients responses? In this case, together two recipients would produce twice the amount of feedback responses observed here than a single recipient.
- On the other hand, it has been observed by Tabensky (2001); Kimbara (2008) that when people collaborate to narrate a story to an audience, they converge their gestures to form shared referent expressions that allows for them to produce a unified and coherent message. Do multiple recipients collaborate to receive a message in the same way by converging their expressions to display a unified front of understanding? How does one recipients feedback displays aid the understanding of a secondary recipient? What is the organisation of recipient collaboration in this way with regards to the issues of epistemic access as highlighted by Heritage (2011)?
- Lastly, Battersby (2011) demonstrated that speaker gaze plays a part when a primary or secondary recipient produces head nodding backchannels. How does this effect multiparty empathetic feedback?

9.8.2 Describing unknown experiences

Currently the descriptions of bodily experiences examined here focus on experiences that are commonly experienced. Most people, if not all, have experienced a toothache or the sensation of yawning. It could be said that this prior knowledge can contribute to the demonstrations of understanding observed here by using their own expression of their own experience to display understanding of the cardholder's description. What is the influence of intersubjectivity when the described experience has not been directly experienced by the recipient? What would be the influence on those recipient responses if the recipient had not had an experience of the same type. Do recipients still produce adapted or elaborations of expressions described by a cardholder from an imagined version of the experience? Alternatively, are embodied resources used differently when an experience cannot be simulated from recall? Would this scenario lead to more clarification sequences similar to the excerpts shown in chapter eight, examples five and six?

9.8.3 Differing motivations to empathise

A limitation of analysing corpora collected in the lab is the artificially structured influence the task/activity has on the interaction. Although the findings are still valuable as an interactional pair still participates in the process of interacting there will be differences in the motivations behind an empathetic engagement with each other's descriptions. Further than the standard moral obligation existing between people to empathetically engage with one another, how do people behave differently when there is a more explicit motivation to empathise? For example, in scenarios when there is an affiliative goal? Lakin et al. (2003, 2008) have demonstrated that prosocial factors such as this have an influence on behavioural mimicry, so it would be expected that they would influence empathetic responses observed here. Do responses alter in therapeutic or counselling scenarios, or medical circumstances where there is a premium in showing an understanding of a patient's experience?

9.8.4 Differing people factors

Presently the analysis looks at a small sample of participants from various western cultures, an equal proportion of male and female participants and an equal mix of strangers and friends.

This gives an overview of the general nonverbal patterns these groups use in conversations about bodily experiences. However, the differences between these groups could provide an insight into how empathetic or mimicry behaviour is influenced by social factors:

- Section 3.1.2 shows that there are differences in the levels of infant-carer mimicry between different cultures. The present corpora only includes participants of western cultures. However, Simpson et al. (2014) show that mothers or carers of Japanese and rural African cultures tend to respond to their infants with contiguous acts without imitation, and whose young produce comparatively less imitation. This suggests that imitative behaviour is experienced from a neonatal stage in western communities. What is the influence of this cultural difference on empathetic behaviour found here such as motor mimicry and expressive adaptions and paraphrases? Are mimicry strategies still be found in alternative cultures that do not engage with mimicry behaviour at an early stage?

- Considering that females are shown to detect and produce nonverbal cues more effectively and sensitively than males during interaction (Hall et al., 2009), is there a difference in the organisation of embodied strategy for empathetic responses as observed here between females and males?

- Neuroimaging studies show an enhanced neural activity in the affective component of the pain matrix when people perceive their loved ones in pain than perceiving pain in strangers Cheng et al. (2010), suggesting neural evidence that people feel more empathy for loved ones over strangers. Is an impact of relationship status reflected in the nonverbal behaviour associated with the empathetic responses?

9.9 Conclusion

The problem of intersubjectivity stems from that our experience of the Self and Other is an asymmetrical phenomenon. To the Self we have direct and continuous experience, to the Other we only experience the outward manifestation of their embodiment. How do we understand the experience of the Other from the outside? Phenomenologists such as Husserl and Merleauu-Ponty consider embodied interaction an important component of how we conceptually resolve the problem of Other minds- how we understand others as minded beings. It is posited that

Other minds can be constituted from the experience that we are similarly embodied, in Husserls' notion of 'pairing' Others are experienced as subject to the same causal relations, the same potentials of preception and therefore, each holding a stream of consciousness as we do. The capacity for this understanding is possible through the constitution of intersubjectivity by the reciprocal experience of another's embodied expressions during interaction- this mutual reflection of understanding illuminates one another as an 'object of perception' as we are. So the key problem of concern here is not just how we understand the 'Other' but how we can have a mutual or intersubjective understanding of each other's experience? This thesis asks in what way our similar embodiment is significant to how we constitute intersubjective understanding of the experience of the Other? In turn, how do we use our bodies to demonstrate this intersubjective empathetic understanding to one another?

Empathy is widely referred to as knowingly but vicariously experiencing the affective state (for instance; emotionally, attitudinally etc) of another. Many models place embodiment as key to our capacity for empathy- these theories stem from the observation that people nonconsciously mimic another's expressions at a pre-attentive level. It is suggested that mimicry of expressions allows for an inner simulation of the affective state associated with an expression due to overlapping motor, sensory and affective neural representations (as mentioned in Section 2.3.1). These overlapping regions are said to be observed as a consequence of the mirror neuron system. When taken together with higher cognitive perspective taking leads to the experience of empathy (Hatfield et al., 1994; Hess and Blairy, 2001; Grèzes and Julie, 2006; Sinigaglia and Carrado, 2011; Gallese, 2013). These models rely ultimately on studies analysing facial mimicry based on responses to photographic or video-based stimulus and not during interaction with another interlocutor. [1]

Similarly for studies of behavioural mimicry as a chameleon effect detailed in Section 3.2.1 taken from research on the perception-behaviour link, proposes that automatic behavioural mimicry of particular behaviours such as face-touch and foot-shaking with confederates is also explained by the perception of an action and the performance of an action sharing a common representational cognitive domain, also said to be observed as a consequence of the mirror neuron system.

[1] Exceptions are found, according to Bourgeois and Hess (2008); Iacobini et al. (2009, 2010), where not all expressions are found to be mimicked, for example, expressions of anger are responded to with neutral expressions, sad expressions are mimicked significantly less than happy expressions.

However, these studies do not align their theory with the simulative accounts of emotional contagion, suggesting instead that the perception-behaviour link is related to the higher cognitive processes of persepctive taking- the top-down component of empathy. Behavioural mimicry argued to be based on a perception-behaviour link highlight the many prosocial effects of mimicking and being mimicked. It is suggested that adopting a similar behaviours and postures when interacting communicate messages indicating an understanding of others' affective states and attitudes, increasing sense of involvedness, togetherness, and consequently, rapport. This tendency is engaged in nonconsciously and automatically[2] (Chartrand and Bargh, 1999; Lakin et al., 2006; Chartrand and van Baaren, 2009; Chartrand and Lakin, 2013).

Is automatic mimicry observed in the present corpus of spontaneous interactions about bodily experiences? Specifically does it extend to gestures that are descriptive or content-specific to the talk? The results from an analysis counting the incidence of descriptive gesture based on five categorical types over the duration of the interaction- (a measure similar to the studies that observe behavioural mimicry studies of foot-shaking / face-touching) were not directly compatible with accounts describing mimicry as 'ubiquitous and engaged in automatically' (Hatfield et al., 1994; Hess and Blairy, 2001; Grèzes and Julie, 2006; Sinigaglia and Carrado, 2011; Gallese, 2013; Chartrand and Lakin, 2013) as cardholders and recipients are found to use different patterns of gesture type and length. Recipients did not appear to mimic the abstract descriptive gestures performed by cardholders. This decrease in engagement with these types of expressions are proposed to explained here as related to higher cognitive processes that govern gesture production; either as the recipient respecting the cardholder's epistemic access, or as due to the inaccessible content of this type of gesture, this type of gesture is engaged with less as would be difficult to establish a common gestural referent of meaning. Another notable difference observed in this analysis was the influence of valence on descriptive gesture. Observed was a longer engagement from recipients for negative experiences. This is on the contrary to some studies of facial mimicry that broadly find less mimicry for negative expressions (Bourgeois and Hess, 2008; Iacobini et al., 2009, 2010). Again, one explanation is that this is also related to higher cognitive processes that govern gesture production- these particular negative experiences included in this corpus provide a repertoire of gestures that are longer in length. In conclusion, it

[2]As significant body of research has also found numerous moderators of this effect, notably all are explained to be related as maintaining the prosocial effect of mimicry.

is suggested that descriptive gestures are not governed by lower cognitive systems of automatic motor responses but higher cognitive systems that are conscious, strategic and interactionally involved. This is in line with studies that find mimicry of descriptive gestures as a strategically utilised display of expressive referents mutually and collaboratively constructed to establish shared concepts.

Although the prevalence of posture congruence has been widely observed in previous studies (Scheflen, 1964, 1972; LaFrance, 1979; Trout and Rosenfeld, 1980; Maurer and Tindall, 1983; Feese et al., 2011; Tia et al., 2011; Hagad et al., 2011), a full bodied and highly sampled analysis of posture similarity found no evidence of postural congruence. The hypothesis that there is a nonconscious and automatic tendency to adopt similar postures, as with behavioural mimicry, is not supported by this study. This is shown where the results show people tend to use certain characteristic movements to describe particular bodily experiences. Additionally, a significant effect of the narrative dialogue structure used in this task on the postures adopted by participants in the cardholder and recipient roles, as opposed to an effect of automatic postural matching. Lastly, this analysis shows physical constitution and habitual patterns of movement have a significant influence differentiating interlocutors posture. Posture appears to be autonomous, or directed by constraints related to the act of communicating rather than automatically matched to an interactional partner.

Here; the models based on facial mimicry that propose simulative underpinnings to empathy, and theories marking the perception-behaviour link as behind occurrences of behavioural mimicry- are not shown to extend their influence to descriptive gesture and bodily posture in this corpus. Observed in the present analyses is that participants behaviour is governed related to higher cognitive processes, that assist in the act of interaction itself. This work aligns itself with the top-down cognitive approach to empathy, that understanding another's experience is constituted via perception of, or in this corpus, the communication of contextual and situational components and a direct understanding of the expressive phenomena displayed within the interaction without the need to internally simulate (see Section 2.2.2, as seen in chapter eight these expressions are carefully negotiated collaboratively and intersubjectively.

This is further demonstrated by the analysis of the organisation of feedback signals, the pattern of feedback responses in these results demonstrates recipients specifically select and place the majority of their feedback signals to function appropriately empathetic to a cardholder's talk,

reciprocally cardholders produce feedback appropriate to their role. Recipients consciously adjust their feedback strategy and use their embodiment in ways appropriate to the material. Clarification sequences show an increase in nonverbal contributions from both conversational parties. This underlines the collaborative nature of conversation, the strategic importance of nonverbal resources for sustaining mutual understanding and the critical role of clarification and repair in working towards intersubjectivity. Nonverbal contributions are primarily and constructively constructed to aid the collaborative, sequential and intersubjective progression of dialogue. This is in line with Garfinkel's, Schegloff's and Schutz' postulation that embodiment is a resource for helping to resolve the practical problems that intersubjectivity present, in the sense that embodiment presents local resources as a means of defending intersubjectivity.

Husserl's concept of 'pairing' relies on the similarities of embodiment a key for our capacity to understand each other's experience through our manifest behaviours. However, it is only through an intersubjective understanding- recognising that others understand our expressions can 'pairing' be fully constituted. Marking communication as a crucial component for empathy and the experience of Other's as minded beings. The examples discussed in chapter eight highlight how demonstrating an understanding of embodied experience is used to reciprocally illuminate similarities in embodiment during interaction. Recipients were found marking similarities by responding to descriptions with expressions that adapt or elaborate on a cardholder's original expressions. As they were not straight imitations- these paraphrases are seen to be sequentially appropriate to demonstrate a richer understanding of the experience being described. Recipients also used non-congruent motor mimicry, an expression that responds in a first person perspective to what is being described, to exploited similarities in embodiment to work towards intersubjectivity. These displays embody the described experience and by doing so demonstrate an understanding of the described experience by enacting appropriate behaviour in that perspective. Lastly, recipients used their similar embodiment to perform a repetition of an expression performed by the cardholder, in order to clarify observed movements in the interest of resolving problems in intersubjectivity (again, in line with Garfinkel's, Schegloff's and Schutz' postulation that embodiment is a resource for helping to resolve the practical problems that intersubjectivity present). Taken together this suggests that people use embodied resources strategically within the context of particular interactional sequences where they provide useful additional resources for sustaining mutual understanding.

The research presented in this thesis has contributed to current views on the role of embodiment to empathy and intersubjectivity. It has introduced new methodology using body movement data in combination with hand-coded annotations and direct observations of audio-visual material to the study of the organisation of embodied resources in spontaneous dyadic interactions designed to promote empathy. This line of study has revealed that automatic mimicry is not as prevalent in spontaneous interaction as the body of research that bases the phenomenon on a perception-behaviour link suggests. On closer inspection of the moment-by-moment organisation of nonverbal behaviour in examples that appear to exhibit behavioural mimicry are intentionally placed and constructed to strategically demonstrate a personal understanding of descriptions of experience. Embodiment is seen to play an essential role in supporting the progression of intersubjectivity. Out of these contributions, this study has introduced an alternative perspective to standard accounts of behavioural mimicry.

Appendices

Queen Mary
University of London

Information sheet

Research study Interaction Corpus: information for participants

We are interested in how people communicate about common experiences. We'll ask you to recall some experiences and talk about them to each other.

The experiences we are interested in are written on sets of cards. You will each be given a stack of these cards and take turns to select one card at a time.

Task:

1. When it is your turn take the top card from your set, read it and discard the card.

2. Try to remember a time you had the experience written on the card and recall the details of the particular sensations you felt at the time.

3. Explain the details of these sensations to your partner so they can understand how this experience felt to you.

Notes:

- You will both have the same cards (they will come up in a different order) but you may each have different sensations to describe for each experience.

- Please feel free to talk or ask questions to each other at any time.

- Try to take no more than 2 minutes for each description.

Bibliography

Alibali, M. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, 44(2):169–188.

Alibali, M. W. (2010). Gesture highlights perceptually present information for speakers. *Gesture*, 10(1):3–28.

Allwood, J. (2003). A study of gestural feedback expressions. *First Nordic Symposium on Multimodal*.

Allwood, J., Cerrato, L., and Jokinen, K. (2007). The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena. *Language resources*.

Allwood, J., Nivre, J., and Ahlsén, E. (1993). On the Semantics and Pragmatics of Linguistic Feedback. *Theoretical Linguistics*, 9(1):1–30.

Anisfeld, M. (1991). Neonatal Imitation. *Developmental Review*, 11:60–97.

Anisfeld, M. (1996). Only Tongue Protrusion Modeling Is Matched by Neonates. *Developmental Review*, 16:149–161.

Ashton-James, C., Baaren, R. V., and Chartrand, T. (2007). Mimicry and me: The impact of mimicry on self-construal. *Social Cognition*, 25(4):518–535.

Athanasiadou, A. (1991). The discourse function of questions. *Pragmatics*, 1(1):107–122.

Atkinson, J. M. and Heritage, J. (1985). *Structures of social action*. Cambridge University Press.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., and Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology*, 16(18):1818–1823.

Bailenson, J. N., Blascovich, J., and Guadagno, R. E. (2008). Self-Representations in Immersive Virtual Environments. *Journal of Applied Social Psychology*, 38(11):2673–2690.

Bailenson, J. N. and Yee, N. (2005). Digital Chameleons Virtual Environments. *Psychological Science*, 16(10):814–819.

Barresi, J. and Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, 19(1):107–122.

Barsalou, L. W., Niedenthal, P. M., Barbey, A. K., and Ruppert, J. A. (2003). Social Embodiment. In Ross, B., editor, *The Psychology of Learning and Motivation*. Elsevier Science.

Bastiaansen, J. A., Thioux, M., and Keysers, C. (2009). Evidence for mirror systems in emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences,*, 364(1528):2391–2404.

Bateson, M. C. (1971). The interpersonal context of infant vocalization. *Quarterly Progress Report of the Research Laboratory of Electronics*, 100:170–176.

Battersby, S. A. (2011). *Moving Together : The organisation of non-verbal cues during multiparty conversation*. Doctoral, Queen Mary University.

Bavelas, J. (2011). The Listener as Addressee in Face-to-face Dialogue. *International Journal of Listening*.

Bavelas, J., Black, A., and Chovil, N. (1988). Form and Function in Motor Mimicry: Topographic Evidence that the Primary Function is Communicative. *Human Communication Research*, 14(3):275–299.

Bavelas, J., Chovil, N., and Lawrie, D. (1992). Interactive gestures. *Discourse Processes*, 15:469–489.

Bavelas, J. B. (2007). *Gesture and the Dynamic Dimension of Language*, chapter Face-to-face Dialogue as a Micro-social Context: The Example of Motor Mimicry, pages 127–146. John Benjamins.

Bavelas, J. B., Black, A., Lemery, C. R., and Mullett, J. (1986). "I Show How You feel": Motor Mimicry as a Communicative Act. *Journal of Personality and Social Psychology*, 50(2):322–329.

Bavelas, J. B., Black, A., Memery, C., and Mullett, J. (1987). *Empathy and its Development*, chapter Motor Mimicry as Primitive Empathy, pages 317–338. Cambridge University Press.

Bavelas, J. B. and Chovil, N. (2000). Visible Acts of Meaning: An Integrated Message Model of Language in Face-to-Face Dialogue. *Journal of Language and Social Psychology*, 19(2):163–194.

Bavelas, J. B., Chovil, N., Coates, L., and Roe, L. (1995). Gestures Specialised for Dialogue. *Personality and Social Psychology Bulletin*, 21:394–405.

Bavelas, J. B., Coates, L., and Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79(6):941–952.

Bavelas, J. B., Coates, L., and Johnson, T. (2006). Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, 52(3):566–580.

Bavelas, J. B., Kenwood, C., Johnson, T., and Phillips, B. (2002). An experimental study of when and how speakers use gestures to communicate. *Gesture*, 2(1):1–17.

Beattie, G. (1988). *The Canderel Guide to Beach Watching*. Rambletree Publishing, Hove.

Beattie, G. and Aboudan, R. (1994). Gestures, pauses and speech: An experimental investigation of the effects of changing social context on their precise temporal relationships. *Semiotica*, 99(3-4):239–272.

Beattie, G. and Coughlan, J. (1999). An experimental investigation of the role of iconic gestures in lexical access using the tip-of-the-tongue phenomenon. *British journal of psychology (London, England : 1953)*, 90 (Pt 1):35–56.

Beattie, G. and Shovelton, H. (2002). What properties of talk are associated with the generation of spontaneous iconic hand gestures? *The British journal of social psychology / the British Psychological Society*, 41(Pt 3):403–17.

Beebe, B., Ph, D., Knoblauch, S., Rustin, J., and Sorter, D. (2003). A Comparison of Meltzoff, Trevarthen, and Stern. *Psychoanalytic Dialogues*, 13(6):809–836.

Bernieri, F. J. and Rosenthal, R. (1991). Interactional Syncrony and Rapport: Measuring Synchrony and in displays devoid of sound and facial effect. *Personality and Social Psychology Bulletin*, 20:303–311.

Bianchi-Berthouze, N., Cairns, P., and Cox, A. (2006). On posture as a modality for expressing and recognizing emotions. *Emotion and HCI*.

Botvinick, M., Jha, A. P., Bylsma, L. M., Fabian, S. A., Solomon, P. E., and Prkachin, K. M. (2005). Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *Neuroimage*, 25(1):312–319.

Bourgeois, P. and Hess, U. (2008). The impact of social context on mimicry. *Biological psychology*, 77(3):343–52.

Brass, M., Bekkering, H., and Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta psychologica,*, 106(1):3–22.

Butterworth, B. and Beattie, G. (1978). *Recent advances in the psychology of language*, chapter Gesture and silence as indicators of planning in speech, pages 347–360. 4b. Springer US.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., and Haggard, P. (2005). Action Observation and Acquired Motor Skills: An fMRI Study with Expert Dancers. *Cerebral Cortex*, 15:1243–1249.

Carpenter, W. B. (1874). *Principles of mental physiology*. John Churchill, London.

Catmur, C., Walsh, V., and Heyes, C. (2007). Sensorimotor Learning Configures the Human Mirror System. *Current Biology*, 17:1527–1531.

Cerrato, L. (2002). Some characteristics of feedback expressions in Swedish. *Proceedings of Fonetik, TMH-QPSR*, 44(1):101–104.

Cerrato, L. (2004). A coding scheme for the annotation of feedback phenomena in conversational speech . In *Proceedings of the LREC Workshop on Models of Human Behaviour for the Specification and Evaluation of Multimodal Input and Output Interfaces*, pages 25–28.

Cerrato, L. (2007). *Investigating Communicative Feedback Phenomena across Languages and Modalities*. Doctoral, Göteborg University.

Chaminade, T., Meltzoff, A. N., and Decety, J. (2002). Does the end justify the means? A PET exploration of the mechanisms involved in human imitation. *NeuroImage*, 15(2):318–28.

Chaminade, T., Zecca, M., Blakemore, S.-J., Takanishi, A., Frith, C. D., Micera, S., Dario, P.,

Rizzolatti, G., Gallese, V., and Umiltà, M. A. (2010). Brain response to a humanoid robot in areas implicated in the perception of human emotional gestures. *PloS one*, 5(7):e11577.

Chartrand, T. and Bargh, J. (1999). The Chameleon Effect: The Perception-Behavior Link and Social Interaction. *Journal of personality and social psychology*, 76(6):893–910.

Chartrand, T. L. and Lakin, J. L. (2013). The Antecedents and Consequences of Human Behavioral Mimicry. *Annual review of psychology*, 64:1–24.

Chartrand, T. L. and van Baaren, R. (2009). Human mimicry. *Advances in experimental social psychology*, 41:219–274.

Cheng, Y., Chen, C., Lin, C. P., Chou, K. H., and Decety, J. (2010). Love hurts: an fmri study. *Neuroimage*, 51(2):923–929.

Chlopan, B. E., McCain, M. L., Carbonell, J. L., and Hagen, R. L. (1985). Empathy: Review of available measures. *Journal of personality and social psychology,*, 48(3):635–653.

Clark, H. (1996). *Using language*. Cambridge University Press.

Clark, H. and Brennen, S. (1991). Grounding in Communication. *Perspectives on Socially Shared Cognition*, 13:222–233.

Clark, H. H. and Schaefer, E. F. (1989). Contributing to discourse. *Cognitive science*, 13(2):259–294.

Condon, W. S. and Sander, L. W. (1974). Neonate movement is synchronized with adult speech: Interactional participation and language acquisition. *Science*, 183(4120):99–101.

Condon, W. S. & Ogston, W. D. (1966). Sound Film Analysis of Nonverbal and Pathological Behaviour Patterns. *The Journal of Nervous and Mental Disease*, 143(1):328–347.

Connolly, W. E. (2012). The Complexity of Intention. *Critical Inquiry*, 37(4):791–798.

Dalton, A. N., Chartrand, T. L., and Finkel, E. J. (2010). The schema-driven chameleon: how mimicry affects executive and self-regulatory resources. *Journal of personality and social psychology,*, 98(4):605–617.

Danziger, N., Prkachin, K. M., and Willer, J.-C. (2006). Is pain the price of empathy? the

perception of others' pain in patients with congenital insensitivity to pain. *Brain*, 129:2494–2507.

De Jaegher, H. (2009). Social understanding through direct perception? Yes, by interacting. *Consciousness and cognition*, 18(2):535–42; discussion 543–50.

de Ruiter, J. P. (1995). Why do people gesture at the telephone? In Biemans M, W. M., editor, *Proceedings of the CLS opening*.

de Ruiter, J. P. (2000). *Language and Gesture*, volume 2, chapter The production of gesture and speech, pages 284–311. Cambridge University Press.

De Silva, P. R. and Bianchi-Berthouze, N. (2004). Measuring posture features saliency in expressing affective states. In *Proceedings of the International Conference on of Intelligent Robots and Systems*, volume 2. IEEE.

de Waal, F. B. M. (2010). *The Age of Empathy*. Harmony Books, New York.

Decety, J. and Chaminade, T. (2003a). Neural correlates of feeling sympathy. *Neuropsychologia*, 41(2):127–38.

Decety, J. and Chaminade, T. (2003b). When the self represents the other: A new cognitive neuroscience view on psychological identification. *Consciousness and Cognition*, 12(4):577–596.

Decety, J. and Sommerville, J. (2003). Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences*, 7(12):527–533.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91:176–180.

Dijkstra, K., Kaschak, M. P., and Zwaan, R. A. (2007). Body posture facilitates retrieval of autobiographical memories. *Cognition,*, 102(1):139–149.

Dinstein, I., Thomas, C., Behrmann, M., and Heeger, D. J. (2008). A mirror up to nature. *Current Biology*, 18(1):13–18.

Duffy, K. A. and Chartrand, T. L. (2015). Mimicry: causes and consequences. *Current Opinion in Behavioral Sciences*, 3:112–116.

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of personality and social psychology*, 23(2):283–292.

Efron, D. (1941). Gesture, race and culture: a tentative study of the spatio-temporal and "linguistic" aspects of the gestural behavior of eastern Jews and southern Italians in New York City living under similar as well as different environmental conditions. *Approaches to Semiotics*, 9:1–216.

Ekman, P. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*.

Enz, S., Zoll, C., Diruf, M., and Spielhagen, C. (2009). Concepts and Evaluation of Psychological Models of Empathy. In *AAMAS 09*. International Foundation for Autonomous Agents and Multiagent Systems.

Etzel, J. A., Gazzola, V., and Keysers, C. (2008). Testing simulation theory with cross-modal multivariate classification of fMRI data. *PloS one,*, 3(11):e3690.

Feese, S., Arnrich, B., Troster, G., Meyer, B., and Jonas, K. (2011). Detecting Posture Mirroring in Social Interactions with Wearable Sensors. In *15th Annual International Symposium on Wearable Computers*, pages 119–120. IEEE.

Ferrari, P. F., Vanderwert, R. E., Paukner, A., Bower, S., Suomi, S. J., and Fox, N. A. (2012). Distinct EEG Amplitude Suppression to Facial Gestures as Evidence for a Mirror Mechanism in Newborn Monkeys. *Journal of Cognitive Neuroscience,*, 24(5):1165–1172.

Finkel, E. J., Campbell, W. K., Brunell, A. B., Dalton, A. N., Scarbeck, S. J., and Chartrand, T. L. (2006). High-maintenance interaction: inefficient social coordination impairs self-regulation. . *Journal of personality and social psychology,*, 91(3):456.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308(5722):662–667.

Fors, K. L. (2011). An investigation of intra-turn pauses in spontaneous speech. Citeseer.

Furuyama, N. (2000). Gestural interaction between the instructor and the learner in origami instruction. In McNeill, D., editor, *Language and Gesture*. University Press, Cambridge.

Gallagher, S. (2001). The practice of mind. Theory, simulation or primary interaction? *Journal of Consciousness Studies, 8*.

Gallagher, S. (2012). Empathy, Simulation, and Narrative. *Science in Context*, 25(03):355–381.

Gallagher, S. and Zahavi, D. (2008). *The Phenomenological Mind: an introduction to philospohy and cognitive science*. Routledge, London.

Gallese, G. R., Fogassi, L., and Vittorio (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Perspectives on Neuroscience*, 2:661–670.

Gallese, V. (2001). The 'Shared Manifold' Hypothesis: From Mirror Neurons To Empathy. *Journal of Consciousness Studies*, 8(5-7):33–50.

Gallese, V. (2008). Empathy, embodied simulation, and the brain: commentary on Aragno and Zepf/Hartmann. *Journal of the American Psychoanalytic Association*, 56(3):769–81; discussion 803, 805–9.

Gallese, V. (2013). Mirror neurons, embodied simulation and a second-person approach to mindreading. *Cortex*, 49(10):2954–2956.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119(2):593–609.

Gallese, V. and Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12):493–501.

Gallese, V. and Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive neuropsychology,*, 22(3-4):455–479.

Garfinkel, H. (1964). Studies of the Routine Grounds of Everyday Activities. *Social Problems*, 11(3):225–250.

Gavrila, D. M. (1999). The Visual Analysis of Human Movement: A Survey. *Computer Vision and Image Understanding*, 73(1):82–98.

Gazzola, V., Aziz-Zadeh, L., and Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current biology,*, 16(18):1824–1829.

Gazzola, V. and Keysers, C. (2009). The Observation and Execution of Actions Share Motor

and Somatosensory Voxels in all Tested Subjects: Single-Subject Analyses of Unsmoothed fMRI Data. *Cerebral Cortex*, 19(6):1239–1255.

Gentilucci, M. and Dalla Volta, R. (2007). The motor system and the relationships between speech and gesture. *Gesture*, 7(2):159–177.

Gerwing, J. and Allison, M. (2009). The relationship between verbal and gestural contributions in conversation: A comparison of three methods. *Gesture*, 9(3):312–336.

Goffman, E. (1978). Response Cries. *Language*, 54(4):787–815.

Goldin-Meadow, S. (2005). The two faces of gesture: Language and thought. *Gesture*, 5(1-2):241–257.

Goldman, A. (1989). Interpretation Psychologized. *Mind and Language*, 4(3):161–185.

Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading,*. Oxford University Press, Oxford.

Goodwin, C. (1979). *Everyday language: Studies in ethnomethodology*, chapter The interactive construction of a sentence in natural conversation, pages 97–121. Irvington Publishers., New York.

Goodwin, C. (1986). Between and within: Alternative sequential treatments of continuers and assessments. *Human Studies*, 9(2-3):205–217.

Gopnik, A. (2003). *Chomsky and his critics.* , chapter The theory theory as an alternative to the innateness hypothesis. Oxford: Blackwell Publishers.

Gordon, R. M. (1986). Folk psychology as simulation. *Mind and Language*, 1(2):158–171.

Goubert, L., Craig, K., Vervoort, T., Morley, S., Sullivan, M., Williams, A. d. C., Cano, A., and Crombez, G. (2005). Facing others in pain: the effects of empathy. *Pain*, 118:285–288.

Grèzes, J. D. and Julie (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, 1079:4–14.

Griffin, H., Varni, G., Volpe, G., Lourido, G. T., Mancini, M., and Bianchi-Berthouze, N. (2015). Gesture mimicry in expression of laughter. In *International Conference on Affective Computing and Intelligent Interaction*.

Guéguen, N., Martin, A., Meineri, S., and Simon, J. (2013). Using mimicry to elicit answers to intimate questions in survey research. *Field methods*, 25(1):47–57.

Gullberg, M. (1998). *Gesture as a Communication Strategy in Second Language Discourse*. PhD thesis, Lund University.

Gullberg, M. (2009). Giving language a hand: gesture as a cue based communicative strategy. *Lund Working Papers in Linguistics*, 44.

Hadar, U., Steiner, T. J., Grant, E. C., and Clifford Rose, F. (1983). Head movement correlates of juncture and stress at sentence level. *Language and Speech*, 26(2):117–129.

Hagad, J. L., Legaspi, R., Numao, M., and Suarez, M. (2011). Predicting levels of rapport in dyadic interactions through automatic detection of posture and posture congruence. In *Third International Conference on Social Computing*, pages 613–616. IEEE.

Haggard, C. H., Bird, G., Johnson, H., and Patrick (2005). Experience modulates automatic imitation. *Cognitive Brain Research*, 22:233–240.

Hall, J. A., Roter, D. L., Blanch, D. C., and Frankel, R. M. (2009). Nonverbal sensitivity in medical students: Implications for clinical interactions. *Journal of general internal medicine*, 24(11):1217–1222.

Harrigan, J. A. (1985). Lintener's Body Movements and Speaking Turns. *Communication Research*, 12(2):233–250.

Hatfield, E., Cacioppo, J., and Rapson, R. (1994). *Emotional Contagion*. University Press, Cambridge.

Healey, P. G., Lavelle, M., Howes, C., Battersby, S., and McCabe, R. (2013). How listeners respond to speaker's troubles. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, pages 2506–2511.

Healey, P. G. T. (2008). *Languages in Flux*, chapter Interactive misalignment: The role of repair in the development of group sub-languages. College Publications.

Healey, P. G. T., Plant, N., Howes, C., and Lavelle, M. (2015). When words fail: Collaborative gestures during clarification dialogues. In *2015 AAAI Spring Symposium Series*.

Healey, P. G. T., Purver, M., and Howes, C. (2010). Structural divergence in dialogue. In

In Proceedings of the Conference on Architectures and Mechanisms for Language Processing. Citeseer.

Heath, C. (1989). Pain Talk: The Expression of Suffering in the Medical Consultation. *Social Psychology Review*, 52(2):113–125.

Heath, C. (2012). *Making Space for Interaction : Architect ' s Design Dialogues Making Space for Interaction : Architect ' s Design Dialogues*. PhD thesis, Queen Mary University of London.

Hebb, D. (1949). *The Organization of Behavior*. Wiley and Sons, New York.

Helt, M. S., Eigsti, I. M., Snyder, P. J., and Fein, D. A. (2010). Contagious yawning in autistic and typical development. *Child development*, 81(5):1620–1631.

Heritage, J. (2011). *The Morality of Knowledge in Conversation*, chapter Territories of knowledge, territories of experience: empathic moments in interaction, pages 159–183. Cambridge University Press.

Hess, U. and Blairy, S. (2001). Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, 40(2):129–41.

Hess, U. and Fischer, A. (2013). Emotional mimicry as social regulation. *Personality and Social Psychology Review*, 17:142–157.

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6(6):242–247.

Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, 5(6):253–262.

Heyes, C. (2010a). Automatic Imitation. *Psychological Bulletin*, 44(3):1–88.

Heyes, C. (2010b). Where do mirror neurons come from? *Neuroscience and Biobehavioral Reviews*, 34(4):575–583.

Heyes, R. P. C., Catmur, C., and Cecilia (2013). Are Automatic Imitation and Spatial Compatibility Meditated by Different Processes. *Cognitive Science*, 37(4):605–630.

Heylen, D. (2006). Perception of non-verbal emotional listener feedback. *Phenomenology and the Cognitive*.

Heylen, D., Bevacqua, E., and Pelachaud, C. (2011). Generating listening behaviour. *Emotion-Oriented*.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21(7):1229–1243.

Holler, J. and Beattie, G. (2003). Pragmatic aspects of representational gestures: Do speakers use them to clarify verbal ambiguity for the listener? *Gesture*, 3(2):127–154.

Holler, J. and Stevens, R. (2007). The Effect of Common Ground on How Speakers Use Gesture and Speech to Represent Size Information. *Journal of Language and Social Psychology*, 26(1):4–27.

Holler, J. and Wilkin, K. (2009). Communicating common ground: How mutually shared knowledge influences speech and gesture in a narrative task. *Language and Cognitive Processes*, 24(2):267–289.

Holler, J. and Wilkin, K. (2011a). An experimental investigation of how addressee feedback affects co-speech gestures accompanying speakers' responses. *Journal of Pragmatics*, 43:3522–3536.

Holler, J. and Wilkin, K. (2011b). Co-Speech Gesture Mimicry in the Process of Collaborative Referring During Face-to-Face Dialogue. *Journal of Nonverbal Behavior*, 35:133–153.

Hommel, B. (2004). Event files: feature binding in and across perception and action. *Trends in Cognitive Sciences*, 8(11).

Hove, M. J. and Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, 27(6):949–961.

Hsee, C. K., Hatfield, E., and Chemtob, C. (1992). Assessment of the Emotional States of Others: Conscious Judgements Versus Emotional Contagion. *Journal of Social and Clinical Psychology*, 11(2):119–128.

Husserl, E. (1969). *Cartesian Meditations: An Introduction to Phenomenology*. M. Nijhoff, The Hague.

Husserl, E. (2006). *the basic problem of phenomenology: from the lectures, winter semester 1910-1911*. Springer, Dordrecht.

Hyden, L.-C. and Peolsson, M. (2002). Pain Gestures: The Orchestration of Speech and Body Gestures. *Health:*, 6(3):325–345.

Iacobini, M., Gonsalves, T., Berthouze, N. B., and Frith, C. (2009). Creating emotional communication with interactive artwork. In *3rd International Conference of Affective Computing and Intelligent Interaction and Workshops*.

Iacobini, M., Gonsalves, T., Bianchi-Berthouze, N., and Frith, C. (2010). Emotional contagion in interactive art. In *Proceedings of the international conference on Kansei engineering and emotion research,*, Paris, France.

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual review of psychology*, 60:653–70.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., and Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, 3(3):e79.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical Mechanisms of Human Imitation. *Science*, 286(5449):2526–2528.

Inzlicht, M., Gutsell, J. N., and Legault, L. (2012). Mimicry reduces racial prejudice. *Journal of Experimental Social Psychology,*, 48(1):361–365.

Jackson, P. L., Brunet, E., Meltzoff, A. N., and Decety, J. (2006a). Empathy examined through the neural mechanisms involved in imagining how i feel versus how you feel pain. *Neuropsychologia*, 44:752–761.

Jackson, P. L., Meltzoff, A. N., and Decety, J. (2005). How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage*, 24(3):771–9.

Jackson, P. L., Rainville, P., and Decety, J. (2006b). To what extent do we share the pain of others? insight from the neural bases of pain empathy. *Pain*, 125:5–9.

Jacob, P. and Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in cognitive sciences*, 9(1):21–5.

Jacobs, N. and Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*, 56:291–303.

Jacobson, A. J. (2007). Empathy, Primitive Reactions and the Modularity of Emotion. *Canadian Journal of Philosophy*, 36(5S):95–113.

James, W. (1890). *Principles of psychology*. Holt, New York.

James, W. T. (1932). A study of the expression of bodily posture. *The Journal of General Psychology*, 7(2):405–437.

Jeannerod, M. (1999a). A dichotomous visual brain. *Psyche*, 5(25).

Jeannerod, M. (1999b). Visuomotor channels: Their integration in goal-directed prehension. *Human Movement Science*, 18:201–218.

Jefferson, G. (1979). *Everyday Language. Studies in Ethnomethodology*, chapter A technique for inviting laughter and its subsequent acceptance declination, pages 79–96. Irvington Publishers.

Johnston, L. (2002). Behavioral mimicry and stigmatization. *Social Cognition*, 20(1):18–35.

Jones, S. S. (1996). Imitation or exploration? Young infants' matching of adults' oral gestures. *Child development*, pages 1952–1969.

Jones, S. S. (2006). Infants learn to imitate by being imitated. In *Proceedings of the 5th International Conference of Development and Learning*.

Kendon, A. (2004). *Gesture: Visible action as utterance*. University Press, Cambridge.

Kessler, K. and Thomson, L. A. (2010). The embodied nature of spatial perspective taking: embodied transformation versus sensorimotor interference. *Cognition*, 114(1):72–88.

Keysers, C. (2009). Mirror neurons. *Current Biology*, 19(21):971–973.

Keysers, C. and Gazzola, V. (2010). Social neuroscience: mirror neurons recorded in humans. *Current Biology*, 20(8):353–354.

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental brain research,*, 153(4):628–636.

Keysers, C. and Perrett, D. (2004). Demystifying social cognition: a Hebbian perspective. *Trends in Cognitive Sciences,*, 8:501–507.

Kimbara, I. (2006). On gestural mimicry. *Gesture*, 6(1):39–61.

Kimbara, I. (2008). Gesture Form Convergence in Joint Description. *Journal of Nonverbal Behavior*, 32(2):123–131.

Kita, S., van Gijn, I., and van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In Wachsmuth, I. and Fröhlich, M., editors, *Gesture and Sign Language in Human Computer Interaction*, volume 1371, pages 23–35. Springer Berlin Heidelberg.

Kleinsmith, A. and Bianchi-Berthouze, N. (2007). Recognizing affective dimensions from body posture. In *In Proceedings of the International Conference on Affective Computing and Intelligent Interaction*, pages 48–58, Berlin Heidelberg. Springer.

Knoblich, G. and Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science*, 12(6):467–472.

Kohler, E., Keysers, C., Umilta, M. A., Fogassi, L., Gallese, V., and Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science,*, 297(5582):846–848.

Kosonogov, V. (2012). Why the Mirror Neurons Cannot Support Action Understanding. *Neurophysiology*, 44(6):499–502.

Krauss, R. and Chen, Y. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? In Zanna, M., editor, *Advances in experimental social psychology*, pages 389–450. Academic Press, San Diego.

Krauss, R. M. (2000). *Language and Gesture*, volume 2, chapter Lexical Gestures and Lexical Access: A Process Model, pages 261–283. Cambridge University Press.

Krumhuber, E. G., Likowski, K. U., and Weyers, P. (2014). Facial mimicry of spontaneous and deliberate duchenne and non-duchenne smiles. *Journal of Nonverbal Behavior*, 38(1):1–11.

Kuhlen, A. and Brennan, S. (2012). Language in dialogue: when confederates might be hazardous to your data. *Psychonomic Bulletin & Review*.

Kupetz, M. (2014). Empathy displays as interactional achievements- Multimodal and sequential aspects. *Journal of Pragmatics*, 61:4–34.

LaFrance, M. (1979). Nonverbal Synchrony and Rapport: Analysis by the Cross-Lag Panel Technique. *Social Psychology Quarterly*, 42(1):66–70.

LaFrance, M. (1982). *Interaction rhythms: Periodicity in communicative behavior*, chapter Posture mirroring and rapport., pages 279–298. Human Sciences Press, New York.

Lakin, J. L. and Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, 14(4):334–339.

Lakin, J. L., Chartrand, T. L., and Arkin, R. M. (2008). I am too just like you: nonconscious behavioral mimicry as an automatic behavioral response to social exclusion. *Psychological Science*, 19(816–22).

Lakin, J. L., Jefferis, V. E., Cheng, C. M., and Chartrand, T. L. (2003). The Chameleon Effect As Social Glue: Evidence For The Evolutionary Significance Of Nonconscious Mimicry. *Journal of Nonverbal Behavior*, 27(3):145–162.

Lakin, T. L. C., Maddux, W. W., and L., J. (2006). *The New Unconscious*, chapter Beyond the Perception-Behavior Link: The Ubiquitous Utility and Motivational Moderators of Nonconscious Mimicry. Oxford University Press.

Lamm, C., Decety, J., and Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *NeuroImage*, 54(3):2492–502.

Leighton, J., Bird, G., Orsini, C., and Heyes, C. (2010). Social attitudes modulate automatic imitation. *Journal of Experimental Social Psychology*, 46(6):905–910.

Likowski, K. U., Mühlberger, A., Seibt, B., Pauli, P., and Weyers, P. (2008). Modulation of facial mimicry by attitudes. *Journal of Experimental Social Psychology*, 44(4):1065–1072.

Likowski, K. U., Weyers, P., Seibt, B., Stöhr, C., Pauli, P., and Mühlberger, A. (2011). Sad and lonely? sad mood suppresses facial mimicry. *Journal of Nonverbal Behavior*, 35(2):101–117.

Lynn, M. and Mynier, K. (1993). Effect of server posture on restaurant tipping. *Journal of Applied Social Psychology*, 23(8):678–685.

Lyons, D. E., Santos, L. R., and Keil, F. C. (2006). Reflections of other minds: how primate social cognition can inform the function of mirror neurons. *Current Opinion in Neurobiology*, 16:230–234.

Maddux, W. W., Mullen, E., and Galinsky, A. D. (2008). Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology*, 44(2):461–468.

Martin, A., Guéguen, N., and Fischer-Lokou, J. (2010). The impact of guilt on mimicry behavior. *Social Behavior and Personality*, 38(7):987–991.

Maurer, R. and Tindall, J. H. (1983). Effect of Postural Congruence on Client's Perception of Counselor Empathy. *Journal of Counseling Psychology*, 30(2):158–163.

McIntosh, D. N. (2006). Spontaneous facial mimicry, liking and emotional contagion. *Polish Psychological Bulletin*, 37(1):31–42.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, 92(3):350–371.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University Of Press, Chicago.

McNeill, D. (2005). *Gesture and thought*. Continuum Press.

McNeill, D., Cassell, J., and Levy, E. T. (1998). Abstract deixis. *Semiotica*, 95(1/2):5–19.

Mehrabian, A. (1969). Significance of posture and position in the communication of attitude and status relationships. *Psychological Bulletin*, 71(5):359–372.

Meltzoff, A. (2002). Imitation as a Mechanism of Social Cognition: Origins of Empathy, Theory of Mind, and the Representation of Action. In *Blackwell handbook of childhood cognitive*. Blackwell Publishers, Oxford.

Meltzoff, A. N. and Brooks, R. (2001). "Like Me" as a Building Block for Understanding Other Minds: Bodily Acts, Attention, and Intention. In Malle, B. F., Moses, L. J., and Baldwin, D. A., editors, *Intentions and Intentionality*, chapter 8, pages 171–191. The MIT Press.

Meltzoff, a. N. and Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates.

Merleau-Ponty, M. (1996). *Phenomenology of perception*. Routledge Classics, London.

Mol, L. and Krahmer, E. (2009). Alignment in Iconic Gestures: Does it make sense? In R.W.Harvey, B.-J. ., editor, *AVSP*, Norwich.

Mol, L., Krahmer, E., Maes, A., and Swerts, M. (2012). Adaptation in gesture: Converging hands or converging minds? *Journal of Memory and Language*, 66(1):249–264.

Mondada, L. (2007). Multimodal resources for turn-taking: pointing and the emergence of possible next speakers. *Discourse Studies*, 9:194–225.

Mondada, L. (2011). Understanding as an embodied , situated and sequential achievement in interaction. *Journal of Pragmatics*, 43(2):542–552.

Morrison, I., Lloyd, D., di Pellegrino, D., and Roberts, N. (2004). vicarious responses to pain in anterior cingulate cortex: Is empathy a multisensory issue? *Cognitive, Affective and Behavioral Neuroscience*, 4(2):270–278.

Morton, A. (1980). *Frames of Mind*. Oxford University Press.

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., and Fried, I. (2010). Single-Neuron Responses in Humans during Execution and Observation of Actions. *Current Biology*, 20(8):750–756.

Nadel, J. and Dumas, G. (2014). The Interacting Body: Intra- and Interindividual Processes During Imitation. *Journal of Cognitive Education and Psychology*, 13(2):163–174.

Neal, D. T. and Chartrand, T. L. (2011). Embodied emotion perception amplifying and dampening facial feedback modulates emotion perception accuracy. *Social Psychological and Personality Science,*, 2(6):673–678.

Newman-Norland, R. D., van Schie, H. T., van Zuijlen, A. M. J., and Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nature neuroscience*, 10(7):817–818.

Niedenthal, P. M. (2007). Embodying emotion. *Science (New York, N.Y.)*, 316(5827):1002–5.

Niedenthal, P. M. and Barsalou, L. W. (2005). Embodiment in Attitudes , Social Perception , and Emotion Silvia Krauth-Gruber and François Ric. *Personality and Social Psychology Review*, 9(3):184–211.

Osaka, N. A. O., Osaka, M., Morishita, M., Kondo, H., and Fukuyama, H. (2004). A word expressing affective pain activates the anterior cingulate cortex in the human brain: an fmri study. *Behavioural Brain Research*, 153:123–127.

Overgaard, S. (2006). The Problem of Other Minds: Wittgenstein's Phenomenological Persective. *Phenomenology and the Cognitive Sciences*, 5:53–73.

Özyürek, A. (2000). *Language and Gesture*, volume 2, chapter The influence of addressee location on spatial language and representational gestures of direction, pages 284–311. Cambridge University Press.

Parrill, F. and Kimbara, I. (2006). Seeing and Hearing Double: The Influence of Mimicry in Speech and Gesture on Observers. *Journal of Nonverbal Behavior*, 30(4):157–166.

Pascolo, P. B. and Budai, R. (2013). Just how consistent is the mirror neuron system paradigm? *Progress in Neuroscience*, 1(1-4):29–43.

Paukner, A., Suomi, S. J., Visalberghi, E., and Ferrari, P. F. (2009). Capuchin monkeys display affiliation toward humans who imitate them. *Science*, 325(5942):880–883.

Pease, A. (1981). *Body Language: How to read others' thoughts by their gestures.* Sheldon Press, London.

Pfurtscheller, G., Neuper, C., Andrew, C., and Edlinger, G. (1997). Foot and hand area mu rhythms. *International Journal of Psychophysiology*, 26(1):121–135.

Pineda, J. E. (2005). The functional significant of mu rythms: Translating "seeing" and "hearing" into "doing.". *Brain Research Reviews, 50, 57–68*, 50:57–68.

Pobric, G. and De c. Hamilton, A. F. D. C. (2006). Action understanding requires the left inferior frontal cortex. *Current Biology*, 16(5):524–529.

Ponari, M., Conson, M., D'Amico, N. P., Grossi, D., and Trojano, L. (2012). Mapping correspondence between facial mimicry and emotion recognition in healthy subjects. *Emotion*, 12(6):1398–1404.

Poppe, R. (2010). A survey on vision-based human action recognition. *Image and Vision Computing*, 28:976–990.

Preston, S. D. and de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *The Behavioral and brain sciences*, 25(1):1–20; discussion 20–71.

Prinz, W. (1990). *A Common Coding Approach to Perception and Action*, pages 167–201. Springer Berlin Heidelberg, Berlin, Heidelberg.

Provine, R. R. (1996). Laughter. *American Scientist*, 84(1):38–45.

Provine, R. R. (2005). Yawning: The yawn is primal, unstoppable and contagious, revealing the evolutionary and neural basis of empathy and unconscious behavior. *American Scientist*, 93(6):532–539.

Psathas, S. (1973). *Phenomenlogical Sociology: Issues and Applications.* Wiley- Interscience Publications, New York.

Purver, M., Ginzburg, J., and Healey, P. (2003). On the means for clarification in dialogue. *New directions in discourse and dialogue*.

Rauzy, R. B., Ferre, G., Blache, P., Espesser, R., and Stephane. (2007). Backchannels revisited from a multimodal perspective. *Auditory-visual Speech Processing*, pages 1–5.

Ray, E. and Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Developmental science*, 14(1):92–105.

Rieser, H. (2011). Gestures indicating dialogue structure. In *Proceedings of SEMdial*, pages 9–18.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2):131–141.

Rosenthal, R. (1963). On the social psychology of the psychological experiment: The experimenter's hypothesis as unintended determinant of experimental results. *American Scientist*, pages 268–283.

Roth, G. and Dicke, U. (2005). Evolution of the brain and intelligence. *Trends in Cognitive Sciences*, 9(5).

Rowbotham, S., Holler, J., Lloyd, D., and Wearden, A. (2011). How Do We Communicate About Pain? A Systematic Analysis of the Semantic Contribution of Co-speech Gestures in Pain-focused Conversations. *Journal of Nonverbal Behavior*, 36(1):1–21.

Rychlowska, M., Cañadas, E., Wood, A., Krumhuber, E. G., Fischer, A., and Niedenthal, P. M. (2014). Blocking mimicry makes true and false smiles look the same. *PLoS One*, 9(3):e90876.

Saarela, M. V. and Hlushchuk, Y. (2007). The compassionate brain: Humans detect intensity of pain from another's face. *Cerebral Cortex*, 17:230–237.

Sato, W., Fujimura, T., Kochiyama, T., and Suzuki, N. (2013). Relationships among facial mimicry, emotional experience, and emotion recognition. *PLoS One*, 8(3):e57889.

Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain: a journal of neurology. 130 (Pt 9): 2452–2461.*, 130(9):2452–2461.

Scheflen, A. E. (1964). The significance of posture on communication systems,. *Psychiatry*, 27:316–331.

Scheflen, A. E. (1972). *Body Language and the Social Order; Communication as Behavioral Control*. Prentice-Hall, Inc., Englewood Cliffs, N. J.

Schegloff, E. A. (1992). Repair after Next Turn : The Last Structurally Provided Defense of in Conversation1 Intersubjectivity. *American Journal of Sociology*, 97(5):1295–1345.

Schegloff, E. A. (1997). Practices and actions: Boundary cases of otherinitiated repair. *Discourse processes,*, 23(3):499–545.

Scheler, M. (1954). *The Nature of Sympathy*. Routledge and Kegan Paul, London.

Schröder, M. (2003). Experimental study of affect bursts. *Speech Communication*.

Schutz, A. (1990). *Collected papers: The problem of social reality*. Springer.

Shockley, K., Baker, A. A., Richardson, M. J., and Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology*, 33(2):201–208.

Shockley, K., Santana, M. V., and Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology*, 29(2):326.

Sikveland, R. and Ogden, R. (2012). Holding gestures accross turns. *Gesture*, 12(2):166–199.

Simpson, E. A., Murray, L., Paukner, A., and Ferrari, P. F. (2014). The mirror neuron system as revealed through neonatal imitation: presence from birth, predictive power and evidence of plasticity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 369.

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neuroscience and Biobehavioral Reviews*, 30:855–863.

Singer, T. and Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, 1156:81–96.

Singer, T., Seymour, B., O'Doherty, J., and Kaube, H. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303:1157–1162.

Sinigaglia, V. G. and Carrado (2011). What is so special about embodied simulation? *Trends in Cognitive Sciences*, 15(11):512–519.

So, W. C., Kita, S., and GoldinMeadow, S. (2009). Using the hands to identify who does what to whom: Gesture and speech go handinhand. *Cognitive Science,*, 33(1):115–125.

Sonnby-Borgström, M. (2003). Automatic mimicry reactions as related to different levels of information processing. *Scandinavian Journal of Psychology*, 43:433–443.

Soto, J. A. and Levenson, R. W. (2009). Emotion recognition across cultures: the influence of ethnicity on empathic accuracy and physiological linkage. *Emotion (Washington, D.C.)*, 9(6):874–84.

Stel, M., Blascovich, J., McCall, C., Mastop, J., Van Baaren, R. B., and Vonk, R. (2010a). Mimicking disliked others: Effects of a priori liking on the mimicryliking link. *European Journal of Social Psychology,*, 40(5):867–880.

Stel, M., Bos, K., and Bal, M. (2012). On Mimicry and the Psychology of the Belief in a Just World: Imitating the Behaviors of Others Reduces the Blaming of Innocent Victims. *Social Justice Research*, 1(25):14–24.

Stel, M., van Baaren, R. B., Blascovich, J., van Dijk, E., McCall, C., Pollmann, M. M., van

Leeuwen, M. L., Mastop, J., and Vonk, R. (2010b). Effects of a priori liking on the elicitation of mimicry. *Experimental Psychology*, 57(6):412.

Stel, M., Van Baaren, R. B., and Vonk, R. (2008). Effects of mimicking: Acting prosocially by being emotionally moved. *European Journal of Social Psychology,*, 38(6):965–976.

Stel, M., van Dijk, E., and Olivier, E. (2009). You want to know the truth? then don't mimic! *Psychological Science*, 20(6):693–699.

Stel, M. and van Knippenberg, A. (2008). The role of facial mimicry in the recognition of affect. *Psychological Science,*, 19(10):984–985.

Stel, M. and Vonk, R. (2010). Mimicry in social interaction: Benefits for mimickers, mimickees, and their interaction. *British Journal of Psychology*, 101:311–323.

Stern, D. (1985). The interpersonal world of the human infant. *New York: Basic*.

Stivers, T. (2008). Research on Language & Social Interaction Stance , Alignment , and Affiliation During Storytelling : When Nodding Is a Token of Affiliation. *Research of Language and Social Interaction*, 41(1):31–57.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E., and Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26):10587–10592.

Streeck, J. and Hartge, U. (1992). *The Contextualization of Language*, chapter Gestures at the Transition Place, pages 135–57. John Benjamins, Amsterdam.

Stubbe, M. (1997). Are you listening ? Cultural influences on the use of supportive verbal feedback in conversation. *Journal of Pragmatics*, 29:257–289.

Stueber, K. (2006). *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*. MIT Press, Cambridge MA.

Stueber, K. (2008). Empathy. In *The Stanford Encyclopedia of Philosophy*.

Tabensky, A. (2001). Gesture and speech rephrasings in conversation. *Gesture*, 1(2):213–235.

Tanner, R. J., Ferraro, R., Chartrand, T. L., Bettman, J. R., and Van Baaren, R. (2008). Of

chameleons and consumption: The impact of mimicry on choice and preferences. *Journal of Consumer Research,*, 34(6):754–766.

Thrasher, M., Van der Zwaag, M. D., Bianchi-Berthouze, N., and Westerink, J. H. (2011). Mood recognition based on upper body posture and movement conference on affectifeatures. In *In proceeding of the International Conference on Affective Computing and Intelligent Interaction*, pages 377–386, Berlin Heidelberg. Springer.

Tia, B., Saimpont, A., Paizis, C., Mourey, F., Fadiga, L., and Pozzo, T. (2011). Does Observation of Postural Imbalance Induce a Postural Reaction? *PLoS One*, 6(3):e17799.

Tiedens, L. Z. and Fragale, A. R. (2003). Power Moves: Complementarity in Dominant and Submissive Nonverbal Behavior. *Journal of Personality and Social Psychology*, 84(3):558–568.

Tranel, D., Kemmerer, D., Adolphs, R., Damasio, H., and Damasio, A. R. (2010). Neural correlates of conceptual knowledge for actions. *Cognitive Neuropsychology*, 20(3):409–432.

Trevarthen, C. (1998). *Intersubjective Communication and Emotion in Early Ontogeny*, chapter The concept and foundations of infant intersubjectivity, pages 15–46. Cambridge University Press.

Trevarthen, C. and Aitken, K. J. (2001). Infant intersubjectivity: research, theory, and clinical applications. *Journal of child psychology and psychiatry, and allied disciplines*, 42(1):3–48.

Trout, D. and Rosenfeld, H. M. (1980). The Effect of Postural Lean and Body Congruence on the Judgment of Psychotherapeutic Rapport. *Journal of Nonverbal Behavior*, 4(3):176–190.

Turella, L., Pierno, A. C., Tubaldi, F., and Castiello, U. (2009). Mirror neurons in humans: Consisting or confounding evidence? *Brain and Language*, 108:10–21.

Umilta, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., and Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. *Neuron*, 31(1):155–165.

Urgesi, C., Candidi, M., Fabbro, F., Romani, M., and Aglioti, S. M. (2006). Motor facilitation during action observation: topographic mapping of the target muscle and influence of the onlooker's posture. *European Journal of Neuroscience,*, 23(9):2522–2530.

van Baaren, R., Janssen, L., Chartrand, T. L., and Dijksterhuis, A. (2009). Where is the love?

The social aspects of mimicry. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1528):2381–9.

van Baaren, R. B., Fockenberg, D. A., Holland, R. W., Janssen, L., and van Knippenberg, A. (2006). The moody chameleon: The effect of mood on non-conscious mimicry. *Social Cognition*, 24(4):426–437.

Van Baaren, R. B., Holland, R. W., Kawakami, K., and Van Knippenberg, A. (2004a). Mimicry and prosocial behavior. *Psychological science*, 15(1):71–74.

Van Baaren, R. B., Horgan, T. G., Chartrand, T. L., and Dijkmans, M. (2004b). The forest, the trees, and the chameleon: context dependence and mimicry. . *Journal of personality and social psychology,*, 86(3):453–459.

van Baaren, R. B., Maddux, W. W., Chartrand, T. L., De Bouter, C., and van Knippenberg, A. (2003). It takes two to mimic: behavioral consequences of self-construals. *Journal of personality and social psychology,*, 84(5):1093–1102.

Van Der Zee, S., Poppe, R., Taylor, P. J., and Anderson, R. (2015). To freeze or not to freeze: A motion-capture approach to detecting deceit. In *Proceedings of the Hawaii International Conference on System Sciences*.

Van Swol, L. M. (2003). The effects of nonverbal mirroring on perceived persuasiveness, agreement with an imitator, and reciprocity in a group discussion. . *Communication Research,*, 30(4):461–480.

Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press, Cambridge MA.

Verberne, F. M., Ham, J., Ponnada, A., and Midden, C. J. (2013). Trusting digital chameleons: The effect of mimicry by a virtual social agent on user trust . In *International Conference on Persuasive Technology*, pages 234–245, Berlin Heidelberg. Springer.

Vrijsen, J. N., Lange, W. G., Dotsch, R., Wigboldus, D. H., and Rinck, M. (2010). How do socially anxious women evaluate mimicry? a virtual reality study. *Cognition and Emotion*, 24(5):840–847.

Wallbott, H. (1991). Recognition of emotion from facial expression via imitation? Some indirect evidence for an old theory. *British Journal of Social Psychology*, 30(3):207–219.

Wang, Y. and Hamilton, A. F. D. C. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Frontiers in human neuroscience*, 6(June):153.

Wang, Y. and Hamilton, A. F. D. C. (2014). Why does gaze enhance mimicry? placing gaze-mimicry effects in relation to other gaze phenomena. *The Quarterly Journal of Experimental Psychology,*, 67(4):747–762.

Ward, N. (1996). Using prosodic clues to decide when to produce back-channel utterances. In *The Fourth International Conference on Spoken Language Processing*, volume 3, pages 1728–1731.

Ward, N. and Tsukahara, W. (2000). Prosodic features which cue back-channel responses in English and Japanese. *Journal of Pragmatics*, 32(8):1177–1207.

Warren, J. E., Sauter, D. a., Eisner, F., Wiland, J., Dresner, M. A., Wise, R. J. S., Rosen, S., and Scott, S. K. (2006). Positive emotions preferentially engage an auditory-motor "mirror" system. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(50):13067–75.

Wiggins, S. (2013). The social life of 'eugh': Disgust as assessment in family mealtimes. *British Journal of Social Psychology*, 52:489–509.

Wilkinson, S. and Kitzinger, C. (2014). Surprise as an Interactional Achievement : Reaction Tokens in Conversation. *Social Psychology*, 69(2):150–182.

Wilson, M. and Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131(3):460–473.

Wittgenstein, L. (1958). *Philosophical Investigations: Translated by G.E.M. Anscombe.* Basil Blackwell, Oxford.

Yabar, Y., Johnston, L., Miles, L., and Peace, V. (2006). Implicit Behavioral Mimicry: Investigating the Impact of Group Membership. *Journal of Nonverbal Behavior*, 30:97–113.

Yang, W., Luo, Z., Chen, I.-m., and Yeo, S. (2010). A Method for Comparing Human Postures from Motion Capture Data. Technical report, Nanyang Technological University, Singapore.

Zahavi, D. (2001). Beyond empathy. Phenomenological approaches to intersubjectivity. In *Between Ourselves : Second-Person Issues in the Study of Consciousness*. Imprint Academic.

Zahavi, D. (2008). Simulation, projection and empathy. *Consciousness and cognition*, 17(2):514–22.

Zahavi, D. (2010). Empathy, Embodiment and Interpersonal Understanding: From Lipps to Schutz. *Inquiry*, 53(3):285–306.

Zellner, B. (1994). *Fundamentals of speech synthesis and speech recognition.*, chapter Pauses and the temporal structure of speech, pages 41–62. John Wiley.

Zhou, J. (2012). The Effects of Reciprocal Imitation on Teacher–Student Relationships and Student Learning Outcomes. *Mind, Brain, and Education*, 6(2):66–73.